

Semi-hidden target recognition in gated viewer images fused with traditional thermal IR images

Master thesis

Menno A. Smeelen
June 2012

Abstract

Nowadays, for the defense and security community, it is of prior importance to classify threats that are merged in a background while at the same time understanding the context of the entire scene. Traditional TV and Infra-Red (IR) cameras allow for an easy context understanding by providing valuable background and scenery information. Unfortunately, they typically do not allow a human observer to detect and classify semi-hidden targets. This study investigates the added value of the combined use of laser range gated viewer (GV) and IR camera to solve semi-hidden target recognition. To this end, an algorithm is developed to fuse GV and IR images based on a weighted averaging technique and employing existing multi-resolution image representation schemes. Our best fusion method for semi-hidden target recognition is selected from all methods considered by using an Image Quality Metric (IQM) combined with an accurate saliency metric. Both metrics are validated using human conspicuity experiments. For very complex scenarios, we additionally designed a background dimming algorithm that dims the scene either entirely or partially based on the context of the scene (contextual) or locally around the threat, while keeping the threat itself undimmed. The optimal combination of fusion method and amount of dimming is determined by means of a second human conspicuity experiment. In a final human experiment, we tested if moving objects influence the preferred amount of dimming. Our work shows that fusing GV into IR scenery images improves the human recognition task on semi-hidden targets. Moreover it demonstrates that a relatively simple pixel-based approach with a PCA-based weighted fusion scheme is the optimal fusion method among those considered. Additional results show that, especially, so-called contextual dimming improves target recognition in very complex scenarios and that moving objects require slightly more dimming in order to obtain the required performance.

Menno A. Smeelen (1197924)
Pattern Recognition and Bioinformatics Group
Faculty of EEMMCS
Delft University of Technology

Members of the MSc thesis committee

Prof.Dr.Ir. M.J.T. Reinders
Dr. M. Loog
Dr. J.A. Redi
Dr. P.B.W. Schwering

Contents

Abstract.....	2
Content	3
 Part 1: Introduction	 5
 Part 2: Article	 7
Abstract.....	7
1 Introduction	7
2 Framework	9
3 Data description.....	11
4 Methodology.....	11
4.1 Fusion algorithm	12
4.2 Dimming algorithm	17
4.3 Image Quality Metric	19
5 Experimental set-up.....	21
5.1 Fusion experiments.....	22
5.2 Dimming experiments.....	23
5.3 Experiments for video stream.....	24
6 Results.....	24
6.1 Fusion results	24
6.2 Background dimming	28
6.3 Dimming in a video-stream.....	30
7 Discussion.....	31
8 Conclusions	33
Acknowledgements.....	33
References	33

Part 3: Supplementary material	41
1 Introduction	41
2 Laser Range Gated Viewer	42
3 Image fusion using Wavelet decomposition.....	43
4 Image decomposition modules.....	44
5 References	46
 Part 4: Report / research log.....	 48
1 Introduction	48
2 Study goal.....	48
3 Research questions	48
4 Study approach	48
5 Verantwoording activiteiten	49
Appendix A: Tabel registration results.....	88
Appendix B: Resultaten registratie scenario 1 en 7.....	89
Appendix C: tabellen met resultaten IQM op beelden gebruikt voor human in the loop validatie en met correlatie.....	91
Appendix D: resultaten IQM op alle scenarios	92
Appendix E: dimming en IQM dimfit.....	97
Appendix F: Overview Matlab files	100

Part 1: Introduction

This part provides a brief introduction to the study that is performed and gives an overview of the documents (parts 1-4) that are provided as well as the relation in between them. The main document for this study is the article itself (part 2), which provides the methodology as well as the important results of the study.

The main focus of the study is the development of a suitable fusion solution for infra red images with laser gated viewer images in order to solve semi-hidden target recognition. The rationale for this study topic originates from defense requirements. The defense department is searching for better recognition methods, because the environment of operation is becoming more complex. The army has to cope with urban and forest environments. Threats nowadays merge in these environments and make use of natural and manmade objects within this environment. This makes recognition hard whereas accurate and validated classification by an operator is required in order to make the correct decision within juridical and ethical boundaries. Traditional camera systems are not able to provide the required information: this is classification of the threat and context understanding in order to recognize the intention of the threat as well. Traditional systems provide the context of the scene. However, when the threat is semi-hidden or camouflaged these systems lack in providing information for classification capability by an operator. New systems are available for complex scenarios, e.g. the Laser Range Gated Viewer (GV) is a system that is especially suitable to provide threat information of semi-hidden targets. Therefore the multi-sensor set-up of a traditional camera and a GV is promising. For defense it is important that such systems provide one clear image in which both target and context are captured in order to limit required operator capacity. Furthermore time-critical responses are required. The problem demands a study for this sensor set-up. Over here, threat classification within the context (threat-context relation) is called threat recognition.

The goal of this study is to find out if semi-hidden target recognition will be improved when adding Laser Range Gated Viewer (GV) in a multi sensor set-up and to define a fusion method for GV images with IR images that improves the recognition task of humans. The following research questions result from this goal:

1. Does a GV improve semi-hidden target recognition by humans in a multi sensor GV+IR set-up?
2. What is an optimal fusion method for GV with IR images in order to improve semi-hidden target recognition by humans?

In order to achieve the goal the following study outline can be summarized:

First registration of the IR and GV images is done. A proper registration technique is selected and adapted for the scenario. Second, after a literature survey, fusion algorithms are designed based on existing image representation techniques. So called fusion rules, which are mathematical operations for actual fusion, have been developed for these algorithms. Third, in order to select the best suitable fusion technique, a so called Image Quality Metric (IQM) is used to select the optimal fusion algorithm from those considered. More advanced saliency metrics are incorporated in this IQM as well. The IQM is

validated with human conspicuity experiments. Next, in order to improve the recognition in very complex scenes, an algorithm for background dimming is designed. A human experiment is used to define the preferred amount of dimming. Also a dim method is developed based on a saliency metric: so called contextual dimming. In order to define the optimal amount of dimming an adapted IQM is designed: Dim Quality Metric (DQM). This DQM is validated with the human experiment. Finally the influence of moving objects on the preferred amount of dimming is investigated with a human experiment. After the experiments it was possible to conclude if adding a GV in the sensor set-up improves semi-hidden target recognition and that it is possible to define a solution for the problem of semi-hidden target recognition.

New elements are: fusion with laser images, a solution for semi-hidden target recognition, validation of quality metrics using human experiments, improving the IQM with better saliency metrics, contextual dimming and investigation of the influence of moving objects on the preferred amount of dimming. As far as we know, no study looked into image fusion of laser images or provided a solution for a specific scenario like semi-hidden target recognition. Most studies investigated or improved a fusion algorithm and compared it to other fusion methods by evaluating it either with simple statistics or with the basic IQM (without advanced saliency metrics).

The article describes the study that has been performed in detail. It provides: a brief introduction, information about the related work, an explanation of the algorithms that have been developed, a description the experiments that have been performed in order to test and validate the algorithms, the results of the experiments and a discussion. Additionally, a document that contains supplementary material to the article is provided (part 3) in order to explain several topics in more detail for the reader who is less familiar with the theoretical background. These topics are: the laser range gated viewer principle, fusion using the wavelet transform decomposition and image decomposition modules that are used in the study. A report / research log is provided as well (part 4), in order to show the daily work that has been performed and resulted in the article in more detail. This report / log contains results that are not included in the article (e.g. the image registration solution).

Part 2: Article

Abstract

Nowadays, for the defense and security community, it is of prior importance to classify threats that are merged in a background while at the same time understanding the context of the entire scene. Traditional TV and Infra-Red (IR) cameras allow for an easy context understanding by providing valuable background and scenery information. Unfortunately, they typically do not allow a human observer to detect and classify semi-hidden targets. This study investigates the added value of the combined use of laser range gated viewer (GV) and IR camera to solve semi-hidden target recognition. To this end, an algorithm is developed to fuse GV and IR images based on a weighted averaging technique and employing existing multi-resolution image representation schemes. Our best fusion method for semi-hidden target recognition is selected from all methods considered by using an Image Quality Metric (IQM) combined with an accurate saliency metric. Both metrics are validated using human conspicuity experiments. For very complex scenarios, we additionally designed a background dimming algorithm that dims the scene either entirely or partially based on the context of the scene (contextual) or locally around the threat, while keeping the threat itself undimmed. The optimal combination of fusion method and amount of dimming is determined by means of a second human conspicuity experiment. In a final human experiment, we tested if moving objects influence the preferred amount of dimming.

Our work shows that fusing GV into IR scenery images improves the human recognition task on semi-hidden targets. Moreover it demonstrates that a relatively simple pixel-based approach with a PCA-based weighted fusion scheme is the optimal fusion method among those considered. Additional results show that, especially, so-called contextual dimming improves target recognition in very complex scenarios and that moving objects require slightly more dimming in order to obtain the required performance.

Keywords: Image fusion, human recognition, Infra-red, Laser Range Gated Viewer, saliency, Image Quality Metrics, conspicuity experiments.

1 Introduction

Visual target classification and recognition play a key role in military operations and during security tasks. The classification has to be accurate and verified in order to make the correct decision within juridical and ethical boundaries or in the most extreme cases to minimize collateral damage. Defense and security have to cope with complex scenes such as urban or forest environment, in which object classification is difficult. In such situations threats make use of the surroundings by hiding or merging while time-critical responses are required. It is then primordial to differentiate between hostile and harmless people and objects. Therefore it is important to understand the context of the scene as well. Throughout this paper the term *object recognition* applies to object classification with context understanding.

In general, traditional TV and infrared (IR) cameras are able to provide a clear context image in which it is possible to differentiate between natural or manmade landscape and living creatures. Problems arise when the threat is merged within the context and is either well camouflaged or semi-hidden (i.e. partly visible). Hence we propose, for such scenarios, to add a Laser Range Gated Viewer (hereafter GV) to traditional cameras. In fact, in a GV laser system the time of flight of a laser pulse is used to set a range gate around the target. Hence, all the information for the visible part of the target will be provided and all information outside the gate will be suppressed. The GV is able to provide with an image in which the semi-hidden object can be classified both by a human operator and by automatic classification algorithms, whereas the traditional cameras provide the context.

The fusion of these images to a single image will allow for the classification of the threat without loss of context. Thus an operator solely focuses on a single image with increased recognition accuracy and reduced reaction time. In this fused image salient details (edges, lines, points, corners etc) must be maximized as the human visual system is sensitive for these details ([1],[2]) which are important for object fixation and object recognition. For context understanding the background also needs to be preserved [3]. As psychophysical experiments demonstrated that evidence of occlusion also helps humans to better understand the object [4], it is important to preserve these details as well.

The goal of this study is to find out if semi-hidden target recognition will be improved when adding a GV in a multi sensor set-up and to define a suitable fusion method for GV images with IR images that improves the recognition task of human operators. Automatic recognition in order to provide Aided Target Recognition (support in the recognition task) is an interesting subject as well. This, however, is outside the scope of our study. Previous studies have solely focused on improving and validating new fusion techniques by comparison with existing methods. To our knowledge no studies investigated a fusion solution for specific scenarios such as semi-hidden target recognition. Furthermore, the advantages of using laser images in the sensor set-up have never been considered. In this study, the added value of defining a fusion method with laser images as well as the comparison for various suitable fusion techniques and solving semi-hidden target recognition is presented. For this study a dataset with both IR and GV images of several different scenarios of the same scene with a human merged within a natural background is used.

The fusion algorithm presented in this study, is build using a simple average weighted fusion scheme as well as existing modules of multi resolution image decompositions. Considering the human visual system the multi resolution representations is considered as highly suitable because they distinctly represent salient details. Indeed, the human visual system is sensitive for salient details. An optimal fusion method should therefore provide with sufficient spectral and spatial details [1]. In practice, the object boundaries (i.e. edges, corners, curves, lines and points) have to be clear, with good contrast and the least possible noise [2], [4], [5], [6], [7], [8]. To achieve good human perception these salient details should be preserved for both the target as well as background objects and should improve human conspicuity [3]. Well designed fusion rules support this, especially when applied on multi-resolution image decompositions.

The metric to define the success of a fusion method is very important in this study which focuses on selection of a suitable fusion method. Ideally a human experiment including the observer's task is most accurate, however, very time consuming and complex. Therefore, in this study, an Image Quality Metric (IQM), validated by a simple human conspicuity experiment, is used to define the optimal fusion method. In this study we focused on the Piella and Heijmans Image Quality Metric [9] which is based on the metric of Wang and Bovik [10] as it considers the human visual system and is especially suited for fused images because it focuses on important details transferred from both input images into the fused image.

In some scenarios the target turned out to be perfectly merged in the background, i.e. the same intensity for target and background even after optimal fusion. For these scenarios we designed a dimming algorithm and an adapted quality metric in order to consider optimal background dimming. Again results are validated with a human conspicuity experiment. In these experiments also a video stream is analyzed in order to investigate the influence of moving objects on the amount of background dimming. Moving objects are very important for context understanding, but can also result in unwanted focus of attention. Besides, moving objects have different effects in the different dimming methods. Based on the results, an approach for multi-sensor semi-hidden target recognition using GV and IR images can be described and it is possible to assess whether GV improves the human recognition task of semi-hidden targets.

The necessary background information, such as a definition of image fusion and the basic principles, and the related work are both provided in section 2. In section 3 the dataset is described. Besides, required pre-conditions of registration and noise reduction are also provided. Section 4 focuses on the algorithm development. First, the methodology and designed fusion rules used for the fusion algorithm and the modules for image decomposition are presented in subsections 4.2. Then the dimming algorithm is introduced in subsection 4.3. The evaluation methodology used to evaluate both the fusion and the dimming is described in subsection 4.4. Afterwards, the experimental set-up is discussed in section 5. The results are presented and discussed in section 6 and 7 respectively. Finally, conclusions are drawn in section 8.

2 Framework

This work adopts the definition of image fusion that Smith and Heather [11] provide: it is 'the intelligent combination of multi-modality sensor imagery for the purpose of providing an enhanced single view of a scene with extended information content'. This means we aim at combining images from different sources into one single image either pixel by pixel or in a region approach, while preserving important details from the input images, suppressing undesirable features and without introducing artifacts or inconsistencies. Lewis et. al. [12] described an approach for image fusion in which they define a distinction of pixel- and region based fusion. We, accordingly adopt this nomenclature. Pixel based fusion considers the entire image and fusion is applied per pixel. Region based fusion is fusion of segmented regions autonomously.

A fusion algorithm for pixel- and region based image fusion basically consist of an *image representation* method and *fusion rules*. An image representation method is either the initial input image or an image decomposition like the wavelet transform or a Laplacian pyramid. A transform method like the wavelet creates a decomposition with a low frequency component containing the image approximation and high frequency components containing the directional details (basic theory for wavelet transform is provided in [13]). Fusion rules are the actual operators to combine pixels or regions in the image representation. Operators are for instance average weighting or selection.

In the past, extended studies on image fusion have been done. A good overview of the basic fusion theory and methods is given by Smith and Heather [11]. Several studies have focused on developing fusion methods and/or comparing results of fusion methods for multi-modal images. In [12] [14] [15] improved methods of the wavelet, which better represent directional information, are used in image fusion, respectively complexe wavelet, curvelet and contourlet. They compare these methods to the basic wavelet decomposition method by using simple statistics, Root Mean Square (RMS) error or image quality metrics. Another method that better represents directional information using Gabor wavelets is given in [16]. In [17] an improved use of wavelet in image fusion is demonstrated by looking at local variance in the decomposition as measure for actual fusion. They compare the result with simple (non) multi-resolution fusion methods and the basic use of wavelet in image fusion by using simple statistics (e.g. mean and entropy). In [18] a multi-spectral segmentation fusion method is proposed in which region based fusion is applied by using a false color method to highlight the important object after segmentation. Evaluation is done in an experiment by human observers and by comparison to average weighted fusion methods. Another region based method is proposed in [19], in which they use an Intensity Hue Saturation (IHS) representation for a color image and fuse the intensity with IR by using a region based approach on a contourlet representation. They compare the result with basic wavelet representation and IHS fusion methods using entropy and an image quality metric. In [20][21] a color mapping of IR is proposed for fusion of IR images with color images. Color mapping is done using color statistics matching with a look-up table created from a reference color image that contains representative colors. They compare results visually with other color fusion methods; whichever gets closer to the initial scene coloring. In [1] a cognitive evaluation method is proposed in which a number of human subjects are asked to segment objects by highlighting the edges in the fused images as well as initial images and joint contour representations are created for each fused image as well as the initial image. They use the precision-recall measure as evaluation criterion for the joint contour representations of average weighted- and multi-resolution fusion schemes. This method is very accurate, however, due to a large number of required subjects and required effort of each subject this method is also very time consuming.

None of the references above investigate a fusion solution for semi-hidden target recognition. Neither do they consider the advantages of combining laser images into the sensor set-up. These topics constitute the added value of this work.

3 Data description

The GV- and IR images that are used in this study were acquired from the Swedish Defense Research Agency (FOI) who recorded the data for several other studies. The dataset consist of IR and GV image sequences from 7 scenarios of the same scene. In each scenario the human that has to be classified is located at a different position in the scene, with different levels of merging within the vegetation and both standing as well as kneeling. The IR images are 768 by 578 pixels (h x v) with a horizontal field of view (HFOV) of 2.4° and the GV images 640 by 477 pixels (h x v) with a HFOV of 3.8 mrad covering a small area in the IR image around the human. Most of the earlier discussed references use TV and/or IR images approximately of the same size. Figure 1 shows the scenario. In this study it is assumed that target detection has been performed by e.g. change detection, movement detection, LIDAR or RADAR.

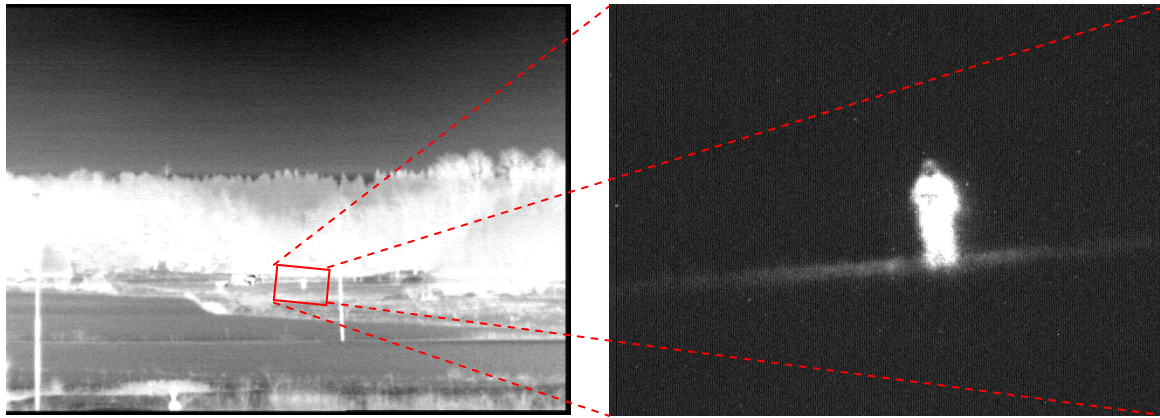


Figure 1: one of the study scenarios provided by Swedish Defense Research Agency (FOI), left IR image and right GV image. The red rectangle in the IR image indicates the area the GV image covers.

Often the input images are poorly aligned, which is also the case in our scenarios (see Figure 1). They need to be aligned in order to avoid incorrect fusion; that a pixel or information will not be fused with its corresponding pixel or information. Therefore, before applying a fusion algorithm to images, image registration of the GV- on the IR images is executed. We used the Elastix toolbox [22] of which rigid and affine transformations with mutual information as similarity metric worked well when combined with a binary mask that indicates the target area. Another characteristic that can negatively influence image fusion is noise. Simple noise removal is done either by Wiener filtering or by thresholding the components of the image transform decomposition.

4 Methodology

In this section the algorithm design is explained. In subsection 4.1, the approach for fusion in our scenario is considered for both pixel- and a region based fusion. The selection of the approach and how to apply it we call *fusion strategy*. Selection of the representation methods is discussed as well and a set of designed fusion rules are proposed for all the representation methods such that important details will be maximized. The background dimming algorithm developed for very complex scenarios in which the object is perfectly merged in the background is described in subsection 4.2. Finally, in subsection 4.3, appropriate image quality metrics (IQM) that estimate the fusion- and dim quality are defined.

4.1 Fusion algorithm

In order to define the fusion strategies (hereafter S) that are suitable for IR and GV images, the scene of our available scenarios is analyzed. Looking into the scene, one can distinguish three separate regions as outlined in Figure 2. The first region outlined in red, from now on called *region I*, is the object to be classified; in this case a human. The most important information for this region is captured in the GV image. The second region outlined in green and outside region *I*, from now on called *region II*, is the area covered by the GV, for which both IR as well as GV contain information. The third region outlined in blue and outside region *II*, from now on called *region III*, is the area only covered by the IR. Therefore, in region *III*, the information from IR is always retained in the algorithm. Within region *I* and *II* a region based- and pixel based fusion approach is followed. We also considered a simple approach of selecting GV information for region *I* and IR for region *II*, which we called *priority fusion*, which is the simplest region based method. However, as explained earlier, the evidence of occlusion often needs to be preserved as well. Therefore, fusion using a variety of methods applies for regions *I* and *II*, either pixel based considering region *I* and *II* as one region or region based considering region *I* and *II* independently. For region based fusion, region *II* is kept IR and fusion only applies for region *I*. The explained strategy S is given by:

$$F = \begin{cases} I: F_i = F_{(IR,GV)_i} \vee F_i = GV_i, & F_i = F_{(IR,GV)_i} \text{ if } II: F_i = F_{(IR,GV)_i} \\ II: F_i = F_{(IR,GV)_i} \vee F_i = IR_i & \\ III: F_i = IR_i & \end{cases}, (1)$$

in which F_i stands for the fused region in which the subscript i stands for the region (*I*, *II* or *III*), $F_{(IR,GV)}$ means fusion of IR and GV in the corresponding region and GV_i and IR_i are respectively GV information or IR information from the corresponding region. F is the final fused image which is a combination of the three fused regions. Note that for pixel based fusion the regions *I* and *II* are one region and $F_i = F_{(IR,GV)}$ for both regions. In the proposed algorithm first region *I* and *II* are cropped from the IR image and up sampled to the same resolution as the GV image, which we call *patch*. Then fusion is applied resulting in a *fused patch* of F_I and F_{II} . The final result F is obtained by down sampling of the fused patch and placing it back in the IR image.

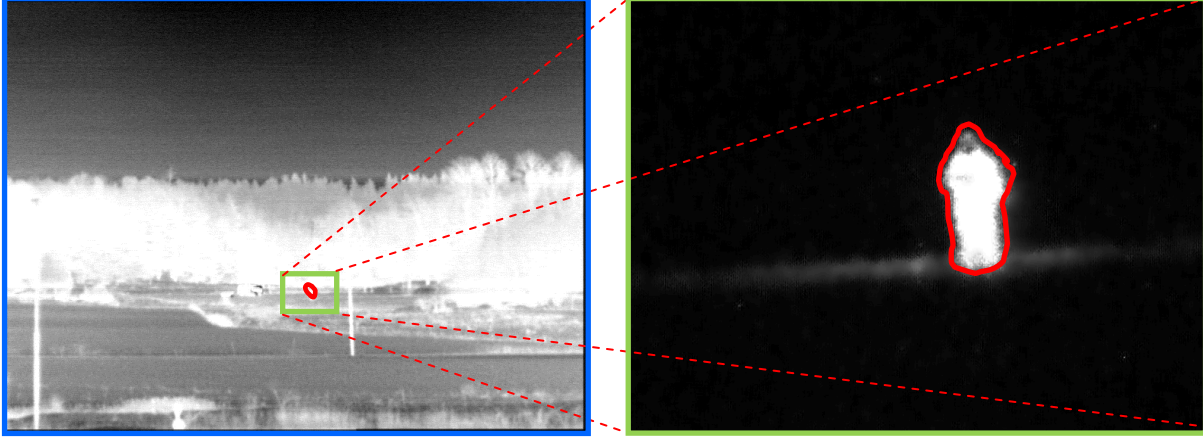


Figure 2: three regions in the scene, again images provided by Swedish Defense Research Agency (FOI), red is region I (the object), green is region II (GV and area in IR covered by GV) and blue is region III (the entire scene in IR)

The following notation is applied for image fusion for both pixel and region based fusion:

$$F_i = \phi(f(I_1), f(I_2))_i, (2)$$

in which I_1 and I_2 are the two input images or regions in which i is the region the fusion applies to (in our case I and/or II), f is the representation method, ϕ the fusion rule and F_i the fusion result. When ϕ results in a weight, the weight is denoted by α . In case f uses an image transformation/decomposition method with inverse transformation/reconstruction the fusion is given by:

$$F_i = f^{-1}(\phi(f(I_1), f(I_2)))_i, (3)$$

Figure 3 shows the flowchart of our algorithm.

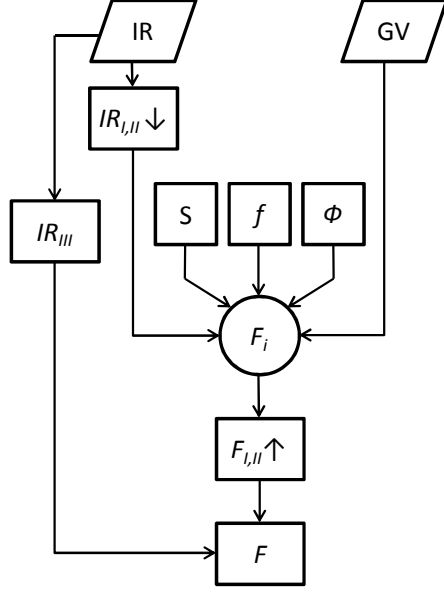


Figure 3: Flowchart of the proposed fusion algorithm. S is the strategy, f is the choice image representation and Φ are the fusion rules. Region I and II of the IR image is down sampled and fused with the GV image according to S and using f and Φ . The fused patch is up sampled and placed back in the initial IR image (it replaces region I and II).

The next step is the selection of image representation or decomposition methods. Weighted average fusion is the simplest fusion method as the fused image is a pixel-by-pixel weighted combination of the two input images. We considered this method because it is simple, fast and easy to implement. This fusion method is given by:

$$F_i = (a \cdot I_1 + (1-a) \cdot I_2)_i, \quad (4)$$

with Φ : $0 \leq a \leq 1$. The weight can be chosen the same for the entire image or different for each pixel. Besides priority fusion and an average weighted fusion scheme multi-resolution methods are selected based on human visual system considerations and used as modules for the algorithm. By selecting transform methods that provide an image decomposition in which directional high frequency components as well as an image approximating low frequency component are included, we will be able to maximize salient details as well as preserve background information by using fusion rules for each component. We selected 3 methods that better represent directional information compared to other multi-resolution decompositions like the basic wavelet transform. The three selected methods are: dual-tree complex wavelet transform (DT-CWT), curvelet transform and contourlet transform. The DT-CWT is developed by N. Kingsbury [23] [24] [25]. Examples for the use in image fusion are provided by Lewis et al. [12] and Nikolov et al. [26]. The curvelet and fast discrete curvelet is proposed by Candès et al. [27] [28] and is especially suitable for objects with smooth edges. Deng et al [14] showed an example of image fusion using curvelet transform. The curvelet transform is available via [29]. The contourlet is proposed by Do and Vetterli [30] and also has flexibility in choosing the number of directions at each level. An example of using the contourlet in region based image fusion of visible and IR images is shown by Ibrahim and Wirth [15]. The contourlet is available via [31]. In order to limit the amount of possible

fusion results, and because the optimal settings of specific representation method is not our goal, the input parameters of the modules (e.g. number of scales/levels) are kept fixed to values that are known to perform well.

Selection of the Φ is an important step in the algorithm design as they define the actual fusion. Both the weights for weighted fusion as well as the fusion of components in the image decomposition modules are defined by Φ . For weighted fusion four different rules are selected. The most simple rule we considered is averaging by weight of $\alpha = 0.5$. Advantage is its simplicity; however, a huge disadvantage is the risk of fading important details. More advanced methods for defining the weight are considered as well. Principle Component Analysis (PCA) is a well known and most often proper functioning method to define the weights [11]. PCA is used to define a weight for the entire image. The eigenvector belonging to highest eigenvalue of the covariance matrix are the weights. PCA gives good results by giving greater weight to the image with more energy. However, in some cases this can also result in a preference for one image and ignoring the other image due to a significant difference in variance. Because of the disadvantages of the above two rules, also two additional rules are considered that define weights per pixel. Both methods define the weight based on the local energy: one by comparing the local variance in a window for each pixel and the other by comparing the local maximum intensity (the sum) of the window. The weight α is defined by $w_1 / (w_1 + w_2)$ in which w_1 and w_2 are the energy values of the windows of respectively image 1 and image 2. The advantage of these rules is that pixel by pixel important information is considered. However, a disadvantage is that it can also result in a speckled pattern. For all four Φ an advantage is that noise can be suppressed; however, a disadvantage is that important features could be suppressed as well.

For the transform/decomposition methods a distinct Φ applies for the low frequency components as well as the high frequency directional components. In general, in most studies, an *averaging* (average weighted) fusion rule is chosen for the low frequency components ($\alpha_{i,j} = 0.5$, i,j being the location in the component) and a *maximum-selection* fusion rule (maximum absolute value at each location in each component) for the directional components. The *maximum-selection* Φ is given by:

$$(f(F)_{m,n})_{i,j} = \max\{abs(f(I_1)_{m,n})_{i,j}, abs(f(I_2)_{m,n})_{i,j}\}, \quad (5)$$

with F being the fused patch (*region I and II*), i,j the location in the component, m,n the component (scale/level and direction) and $\max\{x\}$ a selection (selection in between location i,j from decomposition 1 or 2 that corresponds to the maximum). For the low frequency component an averaging rule makes sure that the approximation of both input images are covered, however, in case of diverging or opposite intensities or different dynamic characteristics of the two modalities this can result in fading of information. This accounts for both the pixel- as well as region based approach. Another possible unwanted effect for pixel based methods is that region II in the GV, that contains little information, will negatively affect the fused patch. This results in a much darker area around the target and can result in an over-fixation of the target by the operator with subsequent loss of context. Therefore, we also considered other Φ :

- *Local-maximum*: a weight per location in the component, based on the sum of all values of a window around this location,

$$\Phi: (a_{i,j})_{m,n} = \frac{(\sum_{w \in W} (f(I_1)_{n,m})|w)}{(\sum_{w \in W} (f(I_1)_{n,m})|w) + (\sum_{w \in W} (f(I_2)_{n,m})|w)}, \quad (6)$$

with w being the window around location i,j and the total set of all windows W . This could result in a background approximation that due to the characteristics of the GV and IR images is closer to the IR image and thus resulting in a smoother fit in the IR context.

- *Local-variance*: a weight per location in the component, based on the variance in a window around this location,

$$\Phi: (a_{i,j})_{m,n} = \frac{(\sigma^2(f(I_1)_{n,m})|w)}{(\sigma^2(f(I_1)_{n,m})|w) + (\sigma^2(f(I_2)_{n,m})|w)}. \quad (7)$$

Although this rule could achieve a similar effect as the local-maximum rule and it also has a lower risk of fading important information (higher weight for locations with more information), there is a risk of speckled pattern in the resulting fused patch.

- *Maximum-local-variance*: selection of the value of the location that has the largest variance in the window around the location,

$$\Phi: (f(F)_{m,n})_{i,j} = \max\{\sigma^2(f(I_1)_{n,m})|w, \sigma^2(f(I_2)_{n,m})|w\}. \quad (8)$$

The same advantage/disadvantage as Local-variance applies here as well.

For the high frequency components with directional information, the maximum-selection rule is in general a good fusion rule as the important salient details that are captured in these components will be preserved. This also has a disadvantage: in case of unwanted features from one modality at a specific location, these will be preserved as well. An example is an edge that is not accurately represented in one image modality due to poor resolution whereas the other image modality represents the edge of the same object more accurate. Both edges will be preserved with the maximum-selection rule resulting in a bad edge artifact in the fused image. Because of this disadvantage we also considered a selection of other Φ :

- *Average*: an average of each component ($a_{i,j} = 0.5$). Averaging in general always has the risk of fading important details, which in this case will affect region II. Still we wanted to see the effect of averaging as somehow we expect all details to be preserved even though it is only slightly.
- *Maximum-average*: selection of the component with the largest average value,

$$\Phi: f(F)_{m,n} = \max\{\mu(f(I_1)_{n,m}), \mu(f(I_2)_{n,m})\}. \quad (9)$$

With this Φ the component with most outstanding details will be selected. However, the details of the component with less outstanding details, which are not present in the corresponding component, will be neglected.

- *Maximum-local-variance*: selection per location based on the largest local variance (variance in the window around the location, see equation 8). The selection is based on local information in the component and not only the location itself. In case of variance this means details that are more pronounced (more contrast) have priority.
- *Local-maximum-selection*: the same as maximum-local-variance, however, in this case the sum of the absolute values of the window rather than the variance,

$$\Phi: (f(F)_{m,n})_{i,j} = \max\{\sum_{w \in W} \text{abs}(f(I_1)_{n,m}) |w, \sum_{w \in W} \text{abs}(f(I_2)_{n,m}) |w\}. \quad (10)$$

The selection is again based on local information in the component and not only the location itself. In case of the sum of absolute values this means details over more locations have priority over details only present at a specific location.

Summarized, Φ based on average values have a risk of fading of important information and Φ based on variance are sensitive to noise. By considering all of the discussed Φ , it is possible to select a suitable Φ for our scenario. In order to limit the amount of results we only considered the described Φ .

4.2 Dimming algorithm

In very complex scenarios, in which after fusion the target still has the same intensity compared to its background, target recognition is still very hard. In these scenarios, background dimming gives a better target background contrast and therefore better target classification. On the other hand, it also decreases the amount of context understanding. There exists a certain point of optimal dimming; a certain amount of dimming for which target classification is improved and context information is still present. In this subsection we focus on the development of suitable dimming algorithms. In the next subsection we show our designed quality metric to define an optimal amount of dimming for our dimming algorithms.

We developed three dimming methods. First of all *global dimming*, which is dimming of the total background by using a dim factor in-between 0 and 1 while the target is remained unchanged. A dim factor that is closer to 0 means more dimming. The second method, or *local dimming*, dimming is only applied to the boundary of the target while using the same dim factor. For both the dimming equation is given by:

$$F_d = d \cdot kF + (1-k)F, \quad (11)$$

in which F is the fused image resulting from equation 1 , d is the dim factor, k a binary mask to indicate the area to be dimmed and F_d the dimmed image. The binary mask k contains 1-s for the total background for global dimming or 1-s for the target boundary for local dimming and 0-s elsewhere.

We also defined a third dimming method that also includes a saliency metric that indicates the areas of the background that contain important context information. This has to be a metric that defines priority regions in an image based on human object fixation. It is important that these metrics include the operator task because importance depends on it. A map that defines and highlights these regions we call a *saliency map* denoted by Sal . The saliency map is used to keep important areas in the background undimmed or less dimmed than unimportant areas. The map reaches a value of 1 for the important areas and zero elsewhere. In this way the background is dimmed with only little loss of context. A smaller dim factor is then possible. We call this *contextual dimming*. Equation 4 is then changed into:

$$F_d = (d + (1-d)Sal) \cdot kF + (1-k)F, (12)$$

in which Sal is a saliency map. For $Sal = 0$ the dim factor is used to dim and elsewhere less dimming applies up to no dimming for $Sal = 1$.

Two ways of creating a more task specific saliency map are distinguished: top-down and bottom-up [32]. Top-down are more complex and more accurate representations of task specific saliency as they highlight the areas that contain important objects with respect to the operator task. Ways to achieve a top-down is either by prioritizing objects with respect to the observers task and train a classifier to perform the task (example methods are Torralba [33] and Navalpakkam and Itti [34]), or by object fixation experiments like an eye movement tracking experiment with the observers as subjects [35]. These top-down methods are very time consuming and complex. Therefore we considered simpler bottom-up methods. Bottom-up methods try to map human priorities by highlighting areas of interest using methods that approximate fixation points or visual attention. The simplest and least performing is variance. The variance shows salient details, however, with no actual approximation of object fixation by humans. Moreover, variance is sensitive to noise. Therefore we considered several more advanced methods suitable for our implementation as well. Several advanced methods are discussed by Toet [32]. We selected the following three methods based on availability and results of comparison with human conspicuity tests in [32]:

- *Frequency Tuned Saliency* or FTS [36]: the Euclidean distance in-between the mean image feature vector and the Gaussian blurred version of the image is used as an saliency metric (representation of DoG), given by

$$Sal(x, y) = \|I_\mu - I_G\|, (13)$$

with (x, y) the pixel position, I_μ the mean image feature vector and I_G the Gaussian blurred image (convolution with kernel of size 5x5). This is performed on CIE Lab color space in order to

achieve saliency as local multi scale color and luminance feature contrast. In our case the image intensity is used as grayscale image and thus the color component is neglected.

- *Harris Points of Interest* or simply Harris [37]: the Harris corner detector is used to indicate points of interest as proposed by M. Loog and F. Lauze [38], since corners indicate presence of objects and interesting features. The Harris corner detector is given by:

$$Sal(x, y) = H(I) = \det T(I) - \kappa \cdot \text{trace}^2 T(I), (14)$$

with H the harris point of interest, I the intensity at its image location and T the structural tensor.

- *Itti-Koch-Niebur* or simply Itti [39]: Itti is a distance weighted multi scale feature dissimilarity map. It uses a pyramidal approach with at each level a feature map for intensity, orientation and color (neglected in our case). These feature maps are combined into a single topographical saliency map. Attention or conspicuous locations are highlighted in order of decreasing saliency.

4.3 Image Quality Metric

The best way to define the optimal fusion method out of the considered methods with respect to recognition by humans is to let humans evaluate the fusion results. A selection of “subjects” needs to perform an objective experiment that is representative for the human recognition task in which the optimal solution can be selected. An Example of such experiments has been presented by Toet et al. [1] [40]. Often a large number of subjects are needed. When the amount of data and fusion methods is large, such an experiment can be very time consuming. Therefore, an image quality algorithm is preferred which is suitable for image fusion and represents the observation task. Wang and Bovik [10] developed a universal objective quality metric, or Image Quality Index (Q_0), which modelles the amount of distortion or common information in between an image and an improved version. The Image Quality Index Q_0 is build up with three components: correlation coefficient, a luminance factor and contrast estimation. This is given by:

$$Q_0 = Q(x, y) = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\mu_x \mu_y}{\mu_x^2 + \mu_y^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}, (15)$$

In which μ_i is the mean, σ_i^2 the variance, σ_{ij} the covariance and x and y the two images. It will give a value in between -1 and 1, 1 being the best quality. It is especially suitable to compare a processed image with a reference image and therefore not directly suitable for fusion. Piella and Heijmans [9] improved the algorithm to be suitable for fusion. They combined $Q(I_1, F)$ with $Q(I_2, F)$ using a weighting system and they perform equation 15 in a sliding window operation such that only local information around the center pixel of the window is taken into account for each corresponding pixel. This Image Quality Metric (IQM) is given by:

$$Q(I_1, I_2, F) = \sum_{w \in W} c(w) (\lambda_w Q_0(I_1, F | w) + (1 - \lambda_w) Q_0(I_2, F | w)), \quad (16)$$

with w being the window of the set of windows W . λ is a weight based on saliency and normally the local variance is taken as saliency metric: $\lambda = \sigma_{I_1}^2 / (\sigma_{I_1}^2 + \sigma_{I_2}^2)$. $c(w)$ is a weighting matrix based on the same saliency metric such that the total sum will be in the range $[-1, 1]$. A weighting system based on saliency is important as it represents the details which are crucial for human recognition. As humans are sensitive for edges and as variance only shows details with no actual approximation of fixation by humans, Piella and Heijmans improved the algorithm even more by incorporating the IQM of the edge representations $Q(I_1', I_2', F')$:

$$Q = Q(I_1, I_2, F) \cdot Q(I_1', I_2', F')^\alpha, \quad (17)$$

α being a parameter that defines the edge contribution. Advantages of the IQM are that it is suitable for all kinds of image fusion and that the priority of the human visual system is incorporated. However, a disadvantage is the lack of a good saliency metric currently used in fusion algorithm evaluation studies. Therefore we considered the saliency maps discussed in subsection 4.2. A saliency map is directly incorporated in the IQM to replace the variance; thus

$\lambda = Sal_1 / (Sal_1 + Sal_2)$. We used both the Wang and Bovik improved by Piella and Heijmans with and without edge representation (see equations 16 and 17) and with variance as well as the three discussed saliency maps for saliency metric.

In order to define an optimal amount of dimming for our dimming algorithms the IQM is adapted. An IQM is designed that is able to define an optimum of the dimming quality. In general, dimming the background decreases the overall quality of the image whereas the quality of the target increases. By multiplying the quality of background with the quality of the target area for each dim factor, a curve arises from which an optimum is obtained. The background quality is defined by the amount of IR context available in the image. Therefore the basic Wang and Bovik algorithm for IR with F_d is used to define the background quality. With this metric more dimming results in lower quality. The image quality of the target area is normally defined by the Piella and Heijmans IQM with both GV as well as IR. However, the IQM of Piella and Heijmans is not suitable because it is already used to define the fusion method and changes in the image will correspond to lower quality. Therefore, another metric is selected for the target quality: target-background contrast. This contrast is defined for a small patch around the human. More dimming corresponds to a larger contrast and thus higher quality. The Michelson-contrast of visibility is selected for target-background contrast, in which the maximum intensity is replaced with the mean intensity of the human and the minimum intensity with the mean intensity of the area around the human. A factor is added to the quality of the target, which functions as a tuning parameter in order to fit the optimum to results of human experiments. The dim quality metric (DQM) then is:

$$Q_d = Q(IR, F_d) \cdot \left(\frac{\overline{F_{d_tgt}} - \overline{F_{d_bg}}}{\overline{F_{d_tgt}} + \overline{F_{d_bg}}} \right)^a, \quad (18)$$

in which $Q(IR, F_d)$ is the image quality according to equation 15 in between the IR image and the dimmed image with dim factor d , $\overline{F_{d_tgt}}$ is the mean intensity of the human, $\overline{F_{d_bg}}$ is the mean intensity of the background and a is the tuning parameter in order to perform fitting to experiments. Moving objects in the image sequence of a scenario do not influence the fusion method; however, we expect that it does influence the optimal amount of dimming. Moving objects are important for context understanding but can also result in object over fixation due to the movement and loss of target fixation. Therefore the influence of moving objects in a video stream on the preferred amount of dimming is considered as well and is included in the experiments of the next section.

5 Experimental set-up

This section gives an overview of the set-up of all experiments and steps taken in order to obtain optimal fusion and optimal dimming from all considered methods for object recognition. The IQM and DQM, as previously described, are used to define the optimal fusion method and optimal dim factor respectively. These metrics are validated by means of human experiments before applying them on the fusion results. Also the influence of the moving objects on the preferred dim factor is tested with a human experiment. The experimental set-up is divided in three sections: fusion algorithm selection, dimming algorithm selection and optimization and tests regarding the influence of moving objects.

First, in subsection 5.1, the required experiments to select the optimal fusion method from the considered methods is described, including validation of the IQM using a human experiment. As explained in subsection 4.2, in very complex scenarios, even after fusion it is still hard to achieve good target recognition for which background dimming is required. Subsection 5.2 explains how a human experiment is used to select the dim method and to fit the DQM for the optimal fusion method and subsequently used to define an optimal dim factor. In subsection 5.3 is explained how a human experiment is used to investigate the influence of moving objects on the preferred optimal dim factor. Figure 4 shows an overview of all steps in the experimental set-up. These steps will be discussed in the corresponding subsections. Also the corresponding subsections which provide the results are shown in the figure.

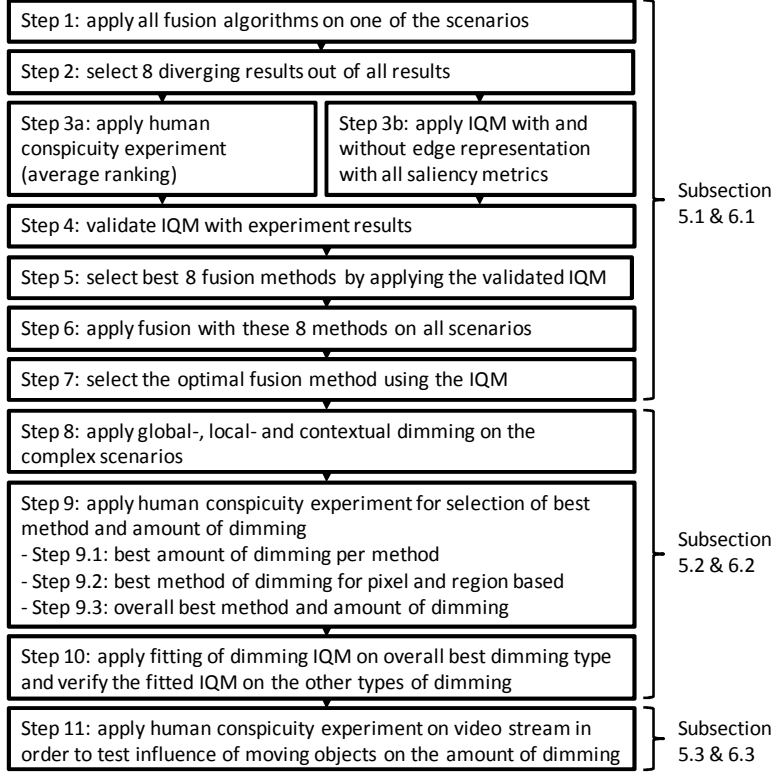


Figure 4: flow chart with steps of the experimental set-up in order to achieve optimal fusion and dimming from all considered methods as well as to test the influence of moving objects on the preferred amount of dimming. The corresponding subsections with experiment descriptions (5.1 – 5.3) as well as corresponding subsections with results (6.1 – 6.3) are shown as well.

5.1 Fusion experiments

The steps 1 to 7 of Figure 4 contain the experiments that provide the optimal fusion method from those considered and also answer if adding a GV to the sensor set-up improves recognition. These experiments also include the validation of the IQM and selection of a saliency metric. The first step is to apply the fusion algorithm with all modules and fusion rules to one of the scenarios. This results in all possible F . All combinations of S , f and Φ give 129 results for F per scenario. 129 results are too many in order to perform a human experiment for validation of the IQM in step 3. Therefore, in step 2, a smaller subset of 8 results is selected (see step 3a).

In step 3 a the human conspicuity experiment is applied on the subset (step 3a) as well as in parallel the IQM is run for all bottom-up saliency metrics (step 3b). The human conspicuity of step 3a needs to be well designed, with an accurate definition of the task of an observer: optimal recognition of the human. To obtain the best results for the task specific recognition, in general, extended experiments as for instance the cognitive model of Toet et. al.[1] , eye movement tracking experiments [35] [41] or conspicuity experiments [40] are good approaches. However, these kinds of experiments are very time consuming, it is complex to include the task and often a large group of subjects are required to obtain useful results. Therefore a simpler and in general well performing ranking experiment is applied. A subset of 8 fused images F and the initial IR image is sufficient for validation of the IQM, as long as they show sufficient visual differences to be able to define a ranking order in the human experiment that is

consistent as well. Therefore, in step 2, a selection of 4 region based as well as 4 pixel based results with visual differences are selected. The experiment of step 3a consists of 14 subjects that rank the images with respect to the recognition task. Ranking is done by giving a lowest score of 1 for the worst image and 9 for the best image. For each image the average ranking is then calculated. An important condition to allow the use of the described simple human experiment is to verify if the ranking is consistent. If the ranking of the subjects show no or only little consistency, the experiment will be unreliable. In parallel, in step 3b the quality is calculated for the fused patches of the same 8 results (before down sampling) and the up sampled patch from IR. It is not useful to run the IQM on the entire scene, because the non fused part (region III) has no influence on the IQM values. The quality is calculated using the IQM both with and without edge representation (see equations 16 and 17) and using the variance as well as the three discussed bottom-up saliency metrics of subsection 4.2.

In order to validate the IQM, in step 4, the average ranking of the human experiment of step 3a and the IQM values of step 3b are compared by performing correlation. Because the human experiment provides a ranking order and the IQM is used to differentiate images based on quality and thus in some way defines a ranking order as well, we also applied the *Spearman* ranking correlation in order to verify the standard *Pearson's* correlation. The IQM and saliency metric combination that has the highest correlation value is selected and used from there on to select the best fusion methods from those considered. In step 5 a set of the best 8 pixel based and region based fusion algorithms are selected from all 129 results using the validated IQM, before performing fusion using this set on the other 6 scenarios in step 6. Finally, by using the validated IQM for all scenarios the overall best performing fusion method can be selected in step 7. This is done by applying a ranking to the IQM values in each scenario followed by calculating an average ranking over all scenarios.

5.2 Dimming experiments

Dimming only applies to the complex scenarios. Step 8, 9 and 10 of the experimental set-up provide the required experiments for these complex scenarios in order to define the optimal dim factor and dimming method from our designed dimming algorithms. First, in step 8, dimming is applied on the best fusion result from step 7 of a complex scenario with all three dimming methods (see equations 11 and 12) and with a dim factor of $d = 0, 0.1, 0.2, \dots, 1$. Next, a second human experiment is done in step 9 in order to select the optimal dim factor and method of dimming. For this second human experiment a set of 16 expert observers were used as subjects. In the first sub step 9.1 of this experiment they were asked to define the optimal dim factor for all three methods of dimming, again optimal being the best recognition. This is asked for an optimal pixel based as well as optimal region based fusion method in order to verify if dimming will influence the preference of fusion. In the second sub step 9.2, they were also asked to select out of these results the optimal dimming method and dim factor for both the pixel- as well as region based method. Finally, in the third sub step 9.3, they were asked to select the overall best dimming method and dim factor combined with the overall best fusion method. By using this method it is possible to achieve a good overview and insight on observer dimming preferences and distribution of the preferred dim factor because the amount of dimming can be different for each fusion- and dimming methods. Finally, after this human experiment, in step 10 the DQM of equation 18 is fitted on the results for all 16 subjects of the overall best dimming method (its distribution) by using

the tuning parameter α . The fitted DQM is in this step also verified for the other dimming methods and it is possible to state if a DQM is effective.

5.3 Experiments for video stream

All previous experiments have been applied on static frames. However, as explained in section 4, moving objects are expected to influence the preferred amount of dimming. Therefore, in step 11, a final experiment has been applied to a scenario with moving objects and only for the overall preferred method of dimming in step 9. For each dim factor $d = 0, 0.1, 0.2, \dots, 1$, a video stream is created and subjects were asked to select the best video stream with optimal object recognition. The dim factor corresponding to the selected video stream is compared with the selected dim factor of step 10.

6 Results

In this section we show the results of the in section 5 described experiments. Subsection 6.1 shows the results to the experiments of step 1 to 7 leading to the selection of the optimal fusion method from those considered. Subsection 6.2 provides the results of the experiments for selection of the optimal dim method and dim factor corresponding to the steps 8 to 10. Finally, in subsection 6.3 the results of the experiment of step 11 regarding the influence of moving objects on the dim factor preferences are provided. In these subsections the order of the steps of section 5 are followed successively.

6.1 Fusion results

Figure 9 shows the IR patch, the GV patch and a selection of 3 region based and 3 pixel based results with visible difference out of the 129 fusion results of step 1. Figure 10 shows the final fused images F for the same results. The shown 6 results are part of the selected subset of 8 results in step 2. The fusion algorithms corresponding to the subset of step 2 which are selected for the human ranking experiment in step 3a are:

1. Pixel based contourlet fusion with average fusion rule ($\alpha = 0.5$) for low frequency components and maximum-selection fusion rule for high frequency components, or PC5
2. Pixel based weighted average fusion with $\alpha = 0.5$, or PW5
3. Pixel based weighted average with weights defined using PCA, or PWP
4. Pixel based contourlet fusion with local-maximum fusion rule for low frequency components and maximum-selection fusion rule for high frequency components, or PCM
5. Region based Priority fusion, or RPF
6. Region based contourlet with local-variance fusion rule for low frequency components and maximum-selection fusion rule for high frequency components, or RCV
7. Region based contourlet with average fusion rule ($\alpha = 0.5$) for low frequency components and maximum-selection fusion rule for high frequency components, or RC5
8. Region based weighted average with weights per pixel defined by local-maximum, or RWM

In Figure 9 and Figure 10 it is clearly visible that fusion improved the initial IR image. It is also apparent that the pixel based fusion methods will result in a darker patch around the human, especially for the weighted fusion using PCA (result 3 or PWP). The local-maximum fusion rule for the low frequency component results in a lighter patch with background intensity that is close to the IR scene intensity (result 4 or PCM), the same as in the case for all region based methods (results 5 -8 or

RPF/RCV/RC5/RWM). Both in the results of the region- as well as the pixel based strategy, it is visible that the multi-resolution methods (results 4 and 7 or PCM/RC5) clearly preserve the directional details, especially on the edge of the human.

Table 1 shows the average ranking results of step 3a for the 8 selected fused images as well as the IR image. A consistency check (see subsection 5.1) showed that all 14 subjects in general gave similar ranking orders.

Table 1: average fused image ranking results of the conspicuity experiment per fused image, 1 being the lowest possible score (worst image) and 9 being the highest possible score (best image). The numbers / acronyms are corresponding to the 8 selected fusion methods listed in the text.

Fused Image	1 PC5	2 PW5	3 PWP	4 PCM	5 RPF	6 RCV	7 RC5	8 RWM	IR
Human Average ranking	7.4	7.4	8.4	2.9	5.6	5.7	4.4	2.0	1.1

As is visible in Table 1, IR has clearly the lowest score. This means that fusion of the GV into the IR scene improves the human recognition task. It can also be observed that the influence of the GV in pixel based fusion has a positive effect on human recognition. When we compared the IQM values for all saliency metric combinations, calculated in step 3b, with the rankings of Table 1 a few interesting facts were apparent. First of all, the human subjects ranked the pixel based methods in general higher than the region based methods, whereas the IQM shows the opposite. An explanation for a higher IQM for region based methods is that the GV has little information outside the object which has a negative influence on the IQM for the pixel based fused images. Though, it is interesting that for the two separate approaches (pixel- and region based) the ranking orders were similar for both subjects as well as IQM. This means the IQM can indeed select the optimal fusion method in the distinctive approaches. A second interesting insight we obtained is the high score of the initial IR image in the IQM with edge representation. This can also be easily explained: most of the edges that do not belong to the target object can only be seen in IR and thus a higher score for IQM on IR edge representation can be expected.

In order to verify that the IQM performs better for distinctive approaches and without IR, the correlation is also determined for three subsets: the total set without IR, pixel based results and region based results. Table 2 shows the respective correlation values that are determined in step 4.

Table 2: correlation values of step 4 in-between the IQM values and the average ranking results of the simple conspicuity experiment, for the IQM with (Q) and without edge representation (Q') and with the different saliency metrics; for all 8 fused images both including the initial IR image (Corr all) and excluding the initial IR image (Corr all without IR) and for pixel based (Corr pixel based) and region based (Corr region based) fusion separately. The closer the correlation value gets to 1 the better the result is.

	Q/locvar	Q/FTS	Q/harris	Q/Itti	Q'/locvar	Q'/FTS	Q'/harris	Q'/Itti
Corr all	0.7499	-0.6943	0.8437	-0.7364	-0.5099	-0.5480	-0.4992	-0.5846
Corr all without IR	0.6296	-0.7437	0.7746	-0.7672	0.4371	0.1864	0.4764	-0.2674
Corr pixel based	0.5680	-0.8195	0.7113	-0.8226	0.8951	0.6527	0.9012	-0.1925
Corr region based	0.9472	0.9495	0.9509	0.0633	-0.8250	-0.8588	-0.8177	-0.8632

The IQM without edge representation and with Harris saliency metric has overall the highest correlation values. Even on the total set of images it reaches a high correlation value. However, on separate pixel- and region based sets the highest IQM values are achieved for respectively IQM with edge representation and Harris saliency metric and IQM without edge representation and Harris saliency metric. In both cases the correlation of the local variance as saliency metric is very close to these results. This means that in our case local variance is performing fairly well. Nevertheless we select the IQM combined with the Harris saliency metric and on pixel and region based separately: pixel based including edge representation and region based without edge representation. Spearman correlation did not change this selection. In step 5 we selected from all 129 fused results the 8 results with the highest IQM value determined with the validated IQM. For region based fusion some results had the same IQM value, therefore the highest 10 values are taken. Step 6 is executed with the fusion algorithms corresponding to these selected results and the IQM is calculated subsequently (step 7). Table 3 and Table 4 respectively show the pixel based and region based rankings of the IQM values including average rankings over all 7 scenarios (denoted by 1 - 7), for all results defined in step 6; 1 being the highest score and 8 the lowest. All rankings of 1 and the overall best (best average ranking) are highlighted in yellow.

Table 3: results of step 7 showing average ranking of the best pixel based fusion methods, for all 7 scenarios, here denoted by 1 to 7. The highest ranked fusion method are highlighted in yellow, both per scenario as well as overall.

Ranking pixel based										
Fusion method / fusion rules (LF, HF or weight)	Scenarios							average rank	rank	
	1	2	3	4	5	6	7			
Contourlet / local-maximum, maximum-average	5	1	5	8	5	7	8	5,57	7	
Contourlet / local-maximum, maximum-local-variance	7	7	7	7	6	8	5	6,71	8	
Curvelet / average, local-maximum-selection	1	5	8	4	7	1	2	4,00	4	
Dtcwt / local-maximum, maximum-average	6	6	6	5	2	2	7	4,86	6	
Dtcwt / local-maximum, local-maximum-selection	8	8	1	3	1	3	3	3,86	3	
WeightedAvg / $\alpha = 0.5$	4	3	2	2	3	6	4	3,43	2	
WeightedAvg / PCA	3	4	3	1	4	4	1	2,86	1	
WeightedAvg / local-maximum	2	2	4	6	8	5	6	4,71	5	

Table 4: results of step 7 showing average ranking of the best region based fusion methods for all 7 scenarios, denoted by 1 to 7. The highest ranked fusion method are highlighted in yellow, both per scenario as well as overall.

Ranking region based									
Fusion method / fusion rules (LF, HF or weight)	Scenarios							average rank	rank
	1	2	3	4	5	6	7		
Contourlet / maximum-local-variance, maximum-selection	8	5	1	1	5	5	1	3,71	5
Contourlet / maximum-local-variance _maxlocabs	8	5	1	1	5	5	1	3,71	5
Contourlet / maximum-local-variance _maxlocvar	8	5	1	1	5	5	1	3,71	5
Curvelet / local-variance, maximum-selection	5	5	8	1	8	8	8	6,14	8
Curvelet / local-variance, local-maximum-selection	5	5	8	1	8	8	8	6,14	8
Curvelet / local-variance, maximum-local-variance	5	5	8	1	8	8	8	6,14	8
Curvelet / maximum-local-variance, maximum-selection	1	1	1	1	1	1	1	1,00	1
Curvelet / maximum-local-variance , local-maximum-selection	1	1	1	1	1	1	1	1,00	1
Curvelet / maximum-local-variance, maximum-local-variance	1	1	1	1	1	1	1	1,00	1
Priority fused	1	1	1	1	1	1	1	1,00	1

The weighted average fusion with PCA scores overall best for the pixel based fusion. It does not give the highest score for all scenarios. But it does show consistency on those scenarios whereas the other fusion algorithms show a widely varying score. An interesting observation is that weighted average with PCA scored also best in the human ranking experiment.

For the region based fusion several fusion algorithms show equal scores. This is due to equal pixel selection in the fusion rules. The inverse transform will give equal results in that case. As visible this is the same as priority fusion: for region I only pixels of the GV are selected. Since fusion using the curvelet transform is less efficient to compute compared to priority fusion, priority fusion is selected as optimal fusion method for region based fusion. Figure 5 gives the images of the two selected optimal fusion methods for one of the scenarios.

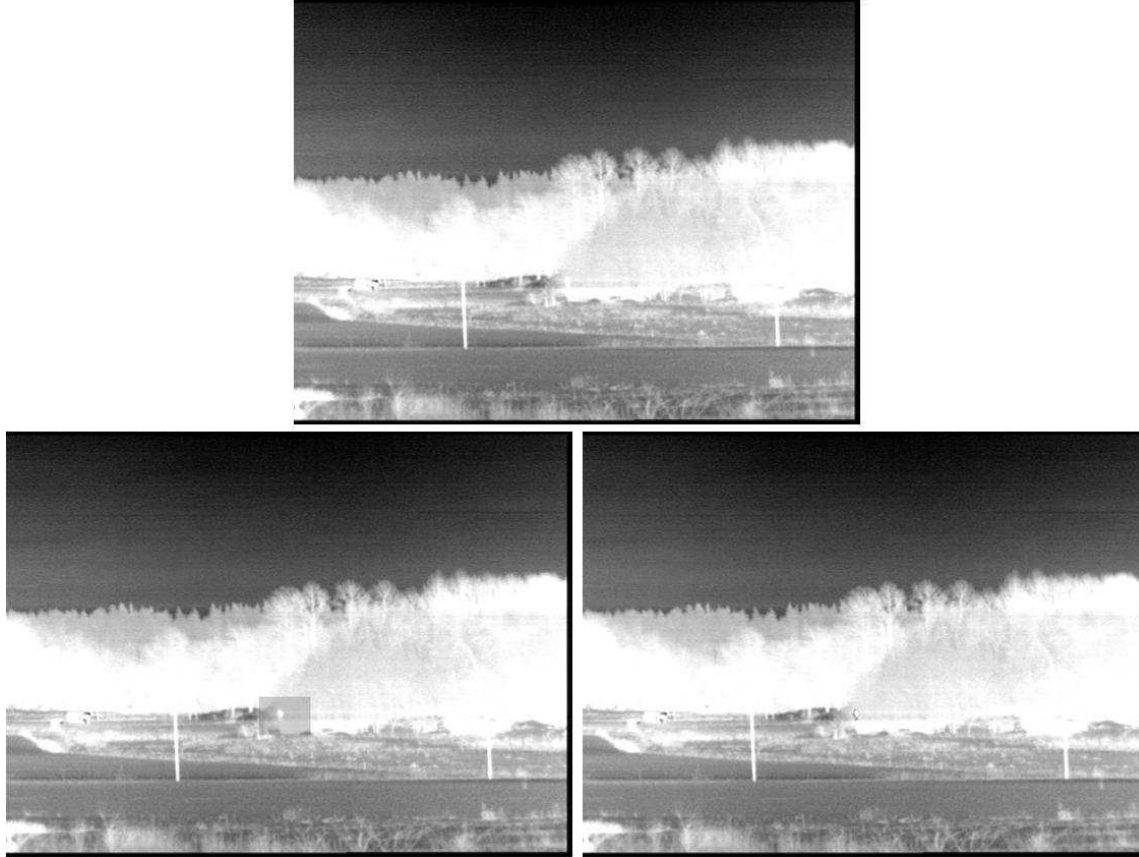


Figure 5: Overall best fusion methods resulting from step 7 for pixel based and region based fusion, top initial IR image (provided by FOI), bottom left pixel based weighted fusion with weights defined with PCA and bottom right region based priority fusion.

It is clear that the scenario of Figure 5 is complex in which the human is very well hidden in between the trees. Although the fusion results in Figure 5 show great improvement of the recognition task, it is still quite difficult to classify the human due to close intensity of object and scene around it. Therefore background dimming is applied to this scenario.

6.2 Background dimming

Figure 6 shows the three ways of dimming as applied in step 8 on the priority fused image and with dim factor of 0.4; clockwise: undimmed, global dimming, local dimming and contextual dimming. We used the Harris saliency map for contextual dimming because it was also selected for the IQM in the first human experiment.

Figure 11 shows the overview of the results from step 9.1, 9.2 and 9.3 on the resulting dimmed images of step 8 for both the selected pixel based and region based fused images. In the two top charts (step 9.1) it can be seen that both contextual as well as global dimming show a distribution with a clear peak. The dim factor that corresponds to this peak is selected for optimal amount of dimming for each method in step 9.1. One can see that the peak for contextual dimming is shifted to the left compared to global dimming, corresponding with slightly more dimming. This emphasizes the use of contextual

dimming. Also it can be observed that local dimming does not show a distribution with a clear peak; dimming improves recognition, but there is not an optimal dim factor. An optimal dim factor can probably be obtained by using a much larger set of subjects. The two charts third from the top (step 9.2) show that contextual dimming is preferred for both pixel- as well as region based fusion. For pixel based fusion a dim factor of 0.6 has a strong preference. For region based fusion a dim factor of 0.6 has an equal preference for both contextual as well as global dimming. However, for several subjects a dim factor of 0.5 also has significant preferences for contextual dimming whereas this is not the case for global dimming. This means that contextual dimming is preferred in general, which is consolidated in the bottom chart that shows the overall preferences (step 9.3). Therefore pixel based fusion combined with contextual dimming is chosen as the preferred method. This is also in agreement with the preference of pixel based fusion in the first experiment. In the top figure it is visible that a dim factor of 0.6 is optimal closely followed by a dim factor of 0.5.

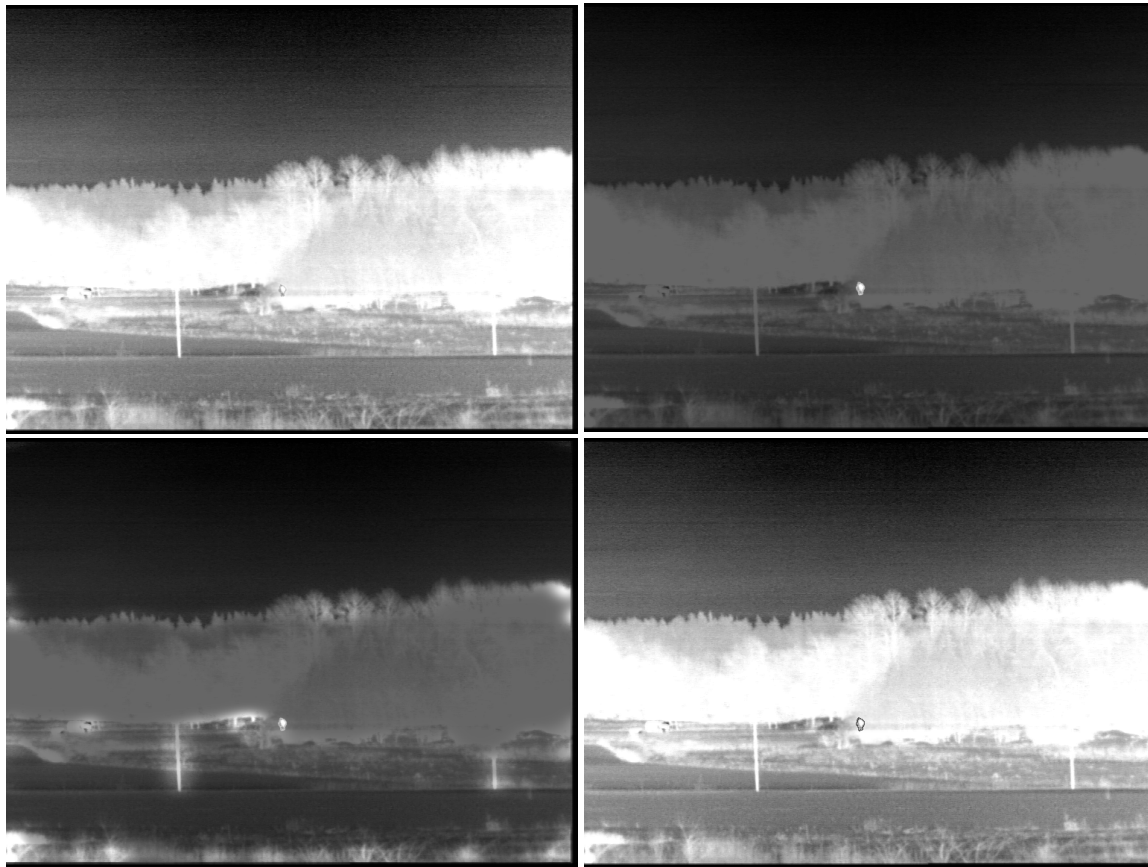


Figure 6: undimmed and dimmed with three types of dimming with dim factor = 0.4, clockwise from top left undimmed, global dimming, local dimming and contextual dimming.

Figure 7 shows the dim quality according to equation 18 against the dim factor for pixel based fusion with contextual dimming and tuning parameter settings of $\alpha = 0.1, 0.2, \dots, 1.5$, in order to apply fitting of the DQM of step 10. A shifting optimum for different tuning parameters is clearly visible. When Figure 7 is compared to the distribution of the human experiment for contextual dimming on pixel based fusion

in the top chart of Figure 11, the best fit is obtained for $a = 0.6$ or $a = 0.7$; the optimum is for the correct dim factor of 0.6 and the next closest dim quality is for a dim factor of 0.5. The optimum is also correct for $a = 0.4, 0.5$ and 0.8 , but then the distribution is not corresponding to the experiment results.

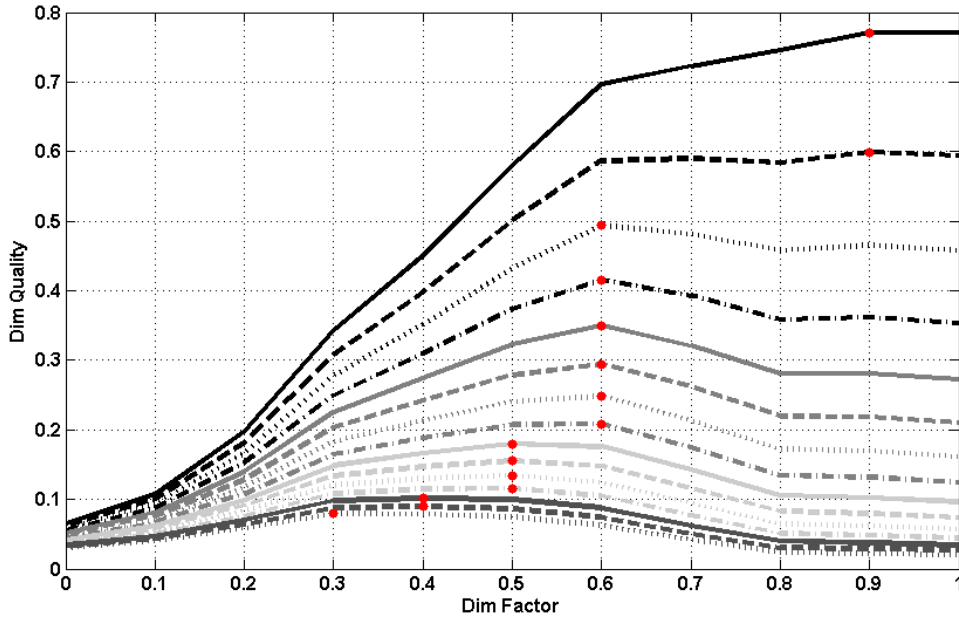


Figure 7: dim quality against dim factor for contextual dimming on pixel based fusion with different steering parameter settings from top to bottom 0.1 to 1.5 in steps of 0.1. Optima are indicated with red dots.

Next, the dim quality is defined with $a = 0.4, 0.5, 0.6, 0.7$ and 0.8 for pixel based fusion with global dimming and region based fusion with both contextual as well as global dimming (see Figure 12). Results are shown in Table 5, in which correct fit is indicated with B (best fit), a correct optimum with V and incorrect fit with X. When observing the results in Table 5, it is apparent that there is no steering parameter setting that provides a best fit for all four combinations, although for $a = 0.4$ and $a = 0.5$ a correct optimum is found for all combinations. Therefore the dim quality can be used to define the optimal amount of dimming; however, it cannot be used for representing a distribution of a human experiment.

Table 5: results of dim quality fitting of steering parameter a , for both pixel based as well as region based fusion and on both contextual as well as global dimming; B = best fit (correct optimum and correct distribution), V = correct optimum and X = incorrect fit.

a	0.4	0.5	0.6	0.7	0.8
Pixelbased/contextual dimming	V	V	B	B	V
Pixelbased/global dimming	V	B	V	X	X
Regionbased/contextual dimming	V	V	B	V	V
Regionbased/global dimming	B	V	X	X	X

6.3 Dimming in a video-stream

The human experiment for the video stream of step 11 is executed with a total of 18 subjects, a slightly larger group than the previous experiments. Figure 8 shows the distribution of all subjects. The peak for

the optimal dim factor moved slightly to a lower value: 0.5. However, the distribution is also wider. In general this means that the observer prefers a slightly larger amount of dimming in a video stream with moving objects compared to a single static frame.

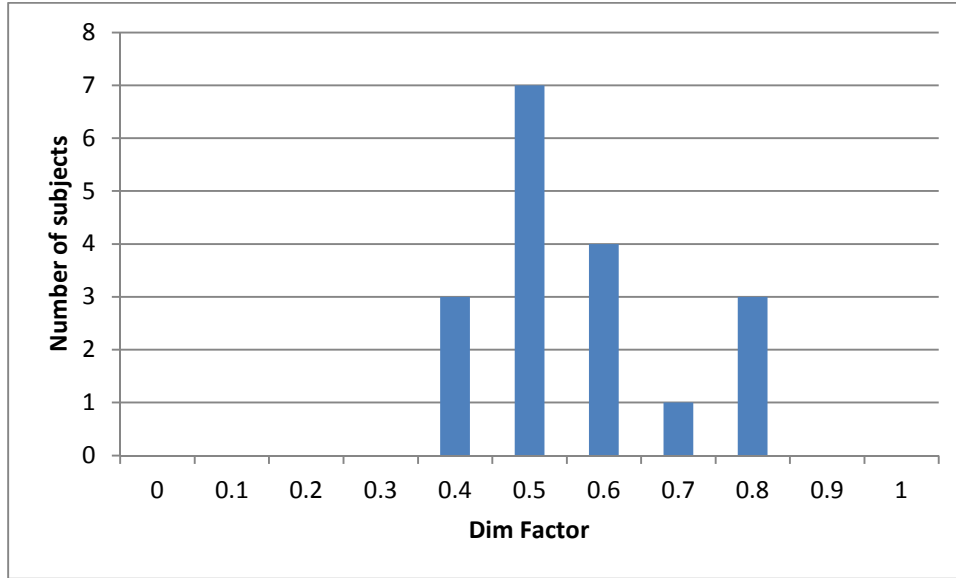


Figure 8: results of the human experiment for influence of moving objects on the optimal dim factor for contextual dimming on pixel based PCA-weighted fusion.

7 Discussion

The experimental results show that fusion of GV with IR improves human recognition and that a simple fusion method like a pixel based approach with a PCA-based weighted fusion scheme is satisfying. Because the GV covers only a small area in the IR image and because it only contains one object, simple fusion methods provide good results. It is obvious that in a scenario of two multi-modal images with a lot of details and different objects a more complex fusion scheme based on the discussed multi resolution methods like the complex wavelet, curvelet or contourlet will probably provide results with better quality. When simple fusion methods provide good results they are preferred because of their simplicity.

In pixel based fusion the IQM showed that multi-resolution methods provide good results and in some scenarios even the best quality. This is due to the fact that IR images contain both background information and details; the more these are preserved in the fused image the better the IQM value is, even though the IR information has some negative influence on the human object itself. The negative influence of IR information on the object and the negative influence of the GV on the background, however, were the reasons to introduce region based fusion in which fusion for the object and background were considered different. As the background is defined by IR, taking IR only seemed the best choice. For the object either GV or a fused object was considered to be the best choice. Although the IQM gave a higher quality for a region based approach, the human experiment showed that the negative influence of GV on the background in pixel based fusion turned out to be a desired influence. It made object fixation easier for a human and the scene is not influenced because the fused patch covers

only a small area in the total IR scene. Thus the darker patch with clearly visible human placed in the lighter IR scene improves the recognition of the human without loss of context. When the entire final image or a much larger patch would be affected, it probably would have a negative influence on the context understanding. One could consider also a smaller area around the human, such that the fused patch will cover even a smaller area in the final fused image. The fact that the IQM showed the opposite with respect to the influence of the patch can be explained by the fact that the IQM looks at the amount of transferred important information and details from both images. As the GV does not contain information and details outside the human object, the quality for this area is determined by the IR image: higher quality for a background that gets closer to the IR image. Both the human experiment as well as the IQM showed that the closer the object gets to the GV object, the better the quality of the fusion result. Because the above described effects, the IQM has to be used for pixel based and region based separately. Considering the above discussion, in future work it might be useful to design and test a region based fusion approach in which the background region in the patch is fused rather than selecting IR.

We considered several bottom-up saliency metrics for incorporating in the IQM. The results show that the Harris algorithm performed best. The results also show that in our scenarios by simply using the local variance as saliency metric the IQM performs fairly well. This is due to the fact that the fused patch only contains one object. We believe that in theory the more objects are present the larger the difference in performance will be for the different saliency metrics because advanced saliency metrics represent important objects better. Therefore it will be interesting to test the IQM with better saliency metrics in other fusion scenarios in order to test this theory. Another improvement will be introducing the top-down saliency metrics in the IQM. This, however, will be a huge effort as complex time consuming experiments need to be performed and, moreover, top-down saliency metric still need to be extensively tested.

We showed by human experiments that in the very complex scenarios, in which the intensity of object and background are the same, background dimming provides an efficient solution for improved recognition. The best dimming method is a contextual dimming method based on the Harris saliency metric (Harris had been selected based on the IQM results). In future the effect of other saliency metrics can be tested, especially in order to incorporate better task dependency. We also showed that an algorithm can be designed that calculates an optimal amount of dimming based on background quality and target background contrast. This works for global as well as contextual dimming. It was not possible to perform fitting of this algorithm for local dimming. A solution to this is to perform a human experiment with a larger set of subjects in order to obtain a clear distribution of dimming preferences. But we have to realize that there might be no clear optimal amount of dimming in local dimming. Our experiment showed that as long as local dimming is performed for a certain dim factor it does not matter how much is dimmed. For local dimming we might speak of a minimum amount of dimming rather than an optimal amount of dimming.

Moving objects in a video stream result in a preference of a slightly more amount of optimal dimming. Therefore, it is valid to say that moving objects do influence the observer's perception. As it was a rather

small experiment on one scenario and only for optimal dimming defined for static frames it might be useful to test and demonstrate the effect of moving objects on other scenarios and types of dimming, especially for scenarios with a larger amount of different types of moving objects, fast and slow moving objects, including the object to be recognized.

8 Conclusions

Both by means of human experiments as well as an Image Quality Metric (IQM) we showed that fusing GV in the IR scenery images with semi-hidden targets improves the human recognition task. We showed that in this scenario there is no need for complex fusion schemes. Furthermore, we demonstrated that in our setting the optimal fusion method is a pixel based approach with a weighted fusion scheme with its weights defined by PCA. The IQM combined with Harris saliency metric turned out to be a proper way to define optimal fusion in our scenarios. It will be interesting to test this IQM on other fusion scenarios and in combination with more bottom-up and top-down task dependent saliency metrics.

We also showed by human experiments, that in very complex scenarios, in which the intensity of object and background are the same, background dimming provides an efficient solution. In our experiment the best dimming method is so called contextual dimming based on the Harris saliency metric, which had been selected by using a new Dimming Quality Metric (DQM) that was fitted on the human experiment results. In general we conclude that for semi-hidden targets adding contextual dimming to a fusion scheme improves the target recognition task.

With a human experiment we showed that in a video stream with moving objects observers prefer more dimming. Therefore, we conclude that moving objects influence the preferred amount of dimming.

Acknowledgements

The authors would like to thank the Swedish Defense Research Agency (FOI), in particular Magnus Elmqvist and Ove Steinvall, for providing us the IR and GV images. We also would like to thank Prof N.G. Kingsbury of the University of Cambridge for providing us the Complex Wavelet implementation.

References

- [1] A. Toet, M.A. Hogervorst, S.G. Nikolov, J.J. Lewis, T.D. Dixon, D.R. Bull, C.N. Canagarajah, "Towards cognitive image fusion," *Elsevier Information Fusion*, vol. 11, pp. 98-113, 2010.
- [2] S. Ullman, "Object recognition and segmentation by a fragmented-based hierarchy," *Elsevier, Trends in cognitive sciences*, vol. 11, no. 2, 2006.
- [3] A.H. Wertheim, "Visual conspicuity: A new simple standard, its reliability, validity and applicability," *Ergonomics*, vol. 53, no. 3, pp. 421-442, 2010.

- [4] R.S. Zemel, M. Behrmann, M.C. Mozer, D. Bavelier, "Experience-dependent perceptual grouping and object-based attention," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 28, no. 1, pp. 202-217, 2002.
- [5] L. Elazary, L. Itti, "Interesting objects are visually salient," *J. Vision*, vol. 8, no. 3, pp. 1-15, 2009.
- [6] J.H. Fecteau, D.P. Munoz, "Saliency, relevance and firing: A priority map for target selection," *Trends in Cognitive Sciences*, vol. 10, no. 8, pp. 382-390, 2006.
- [7] K. Grill-Spector, N. Kanwisher, "Visual recognition: as soon as you know it is there, you know what it is," *Psychological Science*, vol. 16, no. 2, pp. 152-160, 2005.
- [8] L. Itti, C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention," *Vision Research*, vol. 40, no. 10-12, pp. 1489-1506, 2000.
- [9] G. Piella, H. Heijmans, "A new quality metric for image fusion," in *IEEE International Conference on Image Processing*, 2003, pp. 173-176.
- [10] Z. Wang, A.C. Bovik, "A universal image quality index," *IEEE Signal processing letters*, vol. 9, no. 3, pp. 81-84, 2002.
- [11] M.I. Smith, J.P. Heather, "Review of Image Fusion technology in 2005," in *Proc. SPIE 5782, 29 Thermal Image Fusion Applications*, Orlando, FL, USA, 2005.
- [12] J.J. Lewis, R.J. O'Callaghan, S.G. Nikolov, D.R. Bull, "Pixel- and region-based image fusion with complex wavelets," *Elsevier Information Fusion*, vol. 8, pp. 119-130, 2007.
- [13] S.G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693, 1989.
- [14] C. Deng, H. Cao, C. Caob, S. Wang, "Multisensor image fusion using fast discrete curvelet transform," in *Proceedings of SPIE*, vol. 6790 679004-1, 2007.
- [15] S. Ibrahim, M. Wirth, "Visible and IR data fusion technique using the contourlet transform," in *IEEE International Conference on Computational Science and Engineering*, vol. 2, 2009, pp. 42-47.
- [16] Y. Zheng, "An orientation-based fusion algorithm for multisensor image fusion," in *Proceedings of SPIE*, vol. 7710 77100K-1, 2010.
- [17] Z. Guo, J. Yang, "Wavelet transform image fusion based on regional variance," in *Proceedings of SPIE*, vol. 6790 67902Y-1, 2007.
- [18] M. Leviner, M. Maltz, "Multispectral image fusion for target detection," in *Proceedings of SPIE*, vol.

7481 748116-1, 2009.

- [19] T.M. Tu, S.C. Su, H.C. Shyu, P.S. Huang, "A new look at IHS-like fusion methods," *Elsevier Information Fusion*, vol. 2, pp. 177-186, 2001.
- [20] M.A. Hogervorst, A. Toet, "Fast natural colormapping for night-time imagery," *Elsevier Information Fusion*, vol. 11, pp. 69-77, 2010.
- [21] G. Li, S. Xu, X. Zhao, "Fast color-transfer-based image fusion method for merging infrared and visible images," in *Proceedings of SPIE*, vol. 7710 77100S, 2010.
- [22] S. Klein, M. Staring, K. Murphy, M.A. Viergever, J.P.W. Pluim, "elastix: a toolbox for intensity based medical image registration," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 196-205, 2010.
- [23] N.G. Kingsbury, "The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancement," in *Proceedings of EUSIPCO 98*, Rhodes, 1998, pp. 319-322.
- [24] N.G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Journal of Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234-253, May 2001.
- [25] I.W. Selesnick, R.G. Baraniuk, N.G. Kingsbury, "The Dual-Tree Complex Wavelet Transform: A coherent framework for multiscale signal and image processing," *IEEE Signal Processing Magazine*, pp. 123-151, November 2005.
- [26] S.G. Nikolov, P.R. Hill, D.R. Bull, C.N. Canagarajah, "Wavelets for image fusion," in *Wavelets in Signal and Image Analysis, Computational Imaging and Vision Series*. Dordrecht, the Netherlands: Kluwer Academic Publishers, 2001, pp. 213-244.
- [27] E.J. Candès, D.L. Donoho, "Curvelets - A Surprisingly Effective Nonadaptive Representation For Objects with Edges," in *Curves and Surfaces*. Nashville, Tennessee: Vanderbilt University Press, 2000, pp. 105-120.
- [28] E.J. Candès, L. Demanet, D.L. Donoho, L. Ying, "Fast discrete curvelet transform," *Multiscale Modeling and Simulation*, vol. 5, no. 3, pp. 861-899, 2006.
- [29] E.J. Candès, L. Demanet, D.L. Donoho, L. Ying. (2008, April) Curvelet.org, "<http://www.curvelet.org>"
- [30] M.N. Do, M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091-2106, 2005.
- [31] M.N. Do. (2005) Minh N. Do: Software, "<http://www.ifp.illinois.edu/~minhdo/software/>"

- [32] A. Toet, "Computational versus Psychophysical Bottom-up Image Saliency: A Comparative Evaluation Study," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2131-2146, 2011.
- [33] A. Torralba, "Modeling global scene factors in attention," *Journal of Optical Society of America A.*, vol. 20, no. 7, pp. 1407-1418, 2003.
- [34] V. Navalpakkam, L. Itti, "Modeling influence of task on attention," *Elsevier Vision Research*, vol. 45, pp. 205-231, 2005.
- [35] H. Liu, I. Heynderickx, "Visual Attention in Objective Image Quality Assessment: based on Eye-Tracking Data," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, pp. 971-982, 2011.
- [36] R. Achanta, S. Hemami, F. Estrada, S. Süsstrunk, "Frequency-tuned Salient Region Detection," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, 2009.
- [37] C. Harris, M. Stephens, "A combined corner and edge detector," in *Proceedings of Fourth Alvey Vision Conference*, 1988, pp. 147-151.
- [38] M. Loog, F. Lauze, "The Improbability of Harris Interest Points," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 6, pp. 1141-1147, 2010.
- [39] L. Itti, C. Koch, E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254-1259, 1998.
- [40] A. Toet, F.L. Kool, P. Bijl, J.M. Valeton, "Visual conspicuity determines human target acquisition performance," *Optical Engineering*, vol. 37, no. 7, pp. 1969-1975, 1998.
- [41] J. Redi, H. Liu, R. Zunino, I. Heynderickx, "Interactions of visual attention and quality perception," in *IS&T/SPIE Electronic Imaging 2011 and Human Vision and Electronic Imaging XVI*, vol. 7865, 2011.

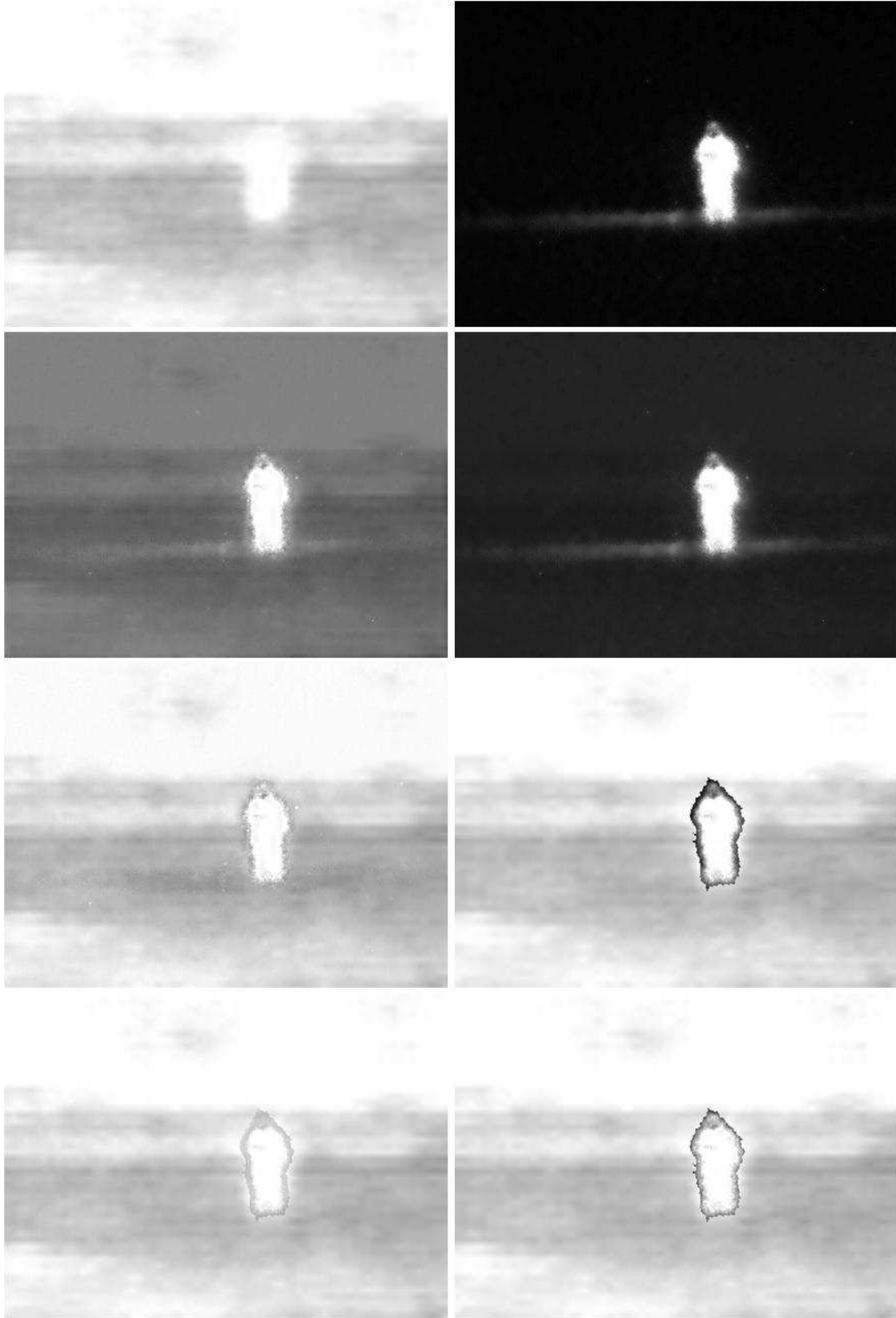


Figure 9: examples of patches (region I and II) for fusion results top to bottom left to right IR, GV (as provided by FOI), 1 (PC5), 3 (PWP), 4 (PCM), 5 (RPF), 8 (RWM), 7 (RC5). The numbers and acronyms are corresponding to the human conspicuity test in section 6.1 and the results in Table 1.

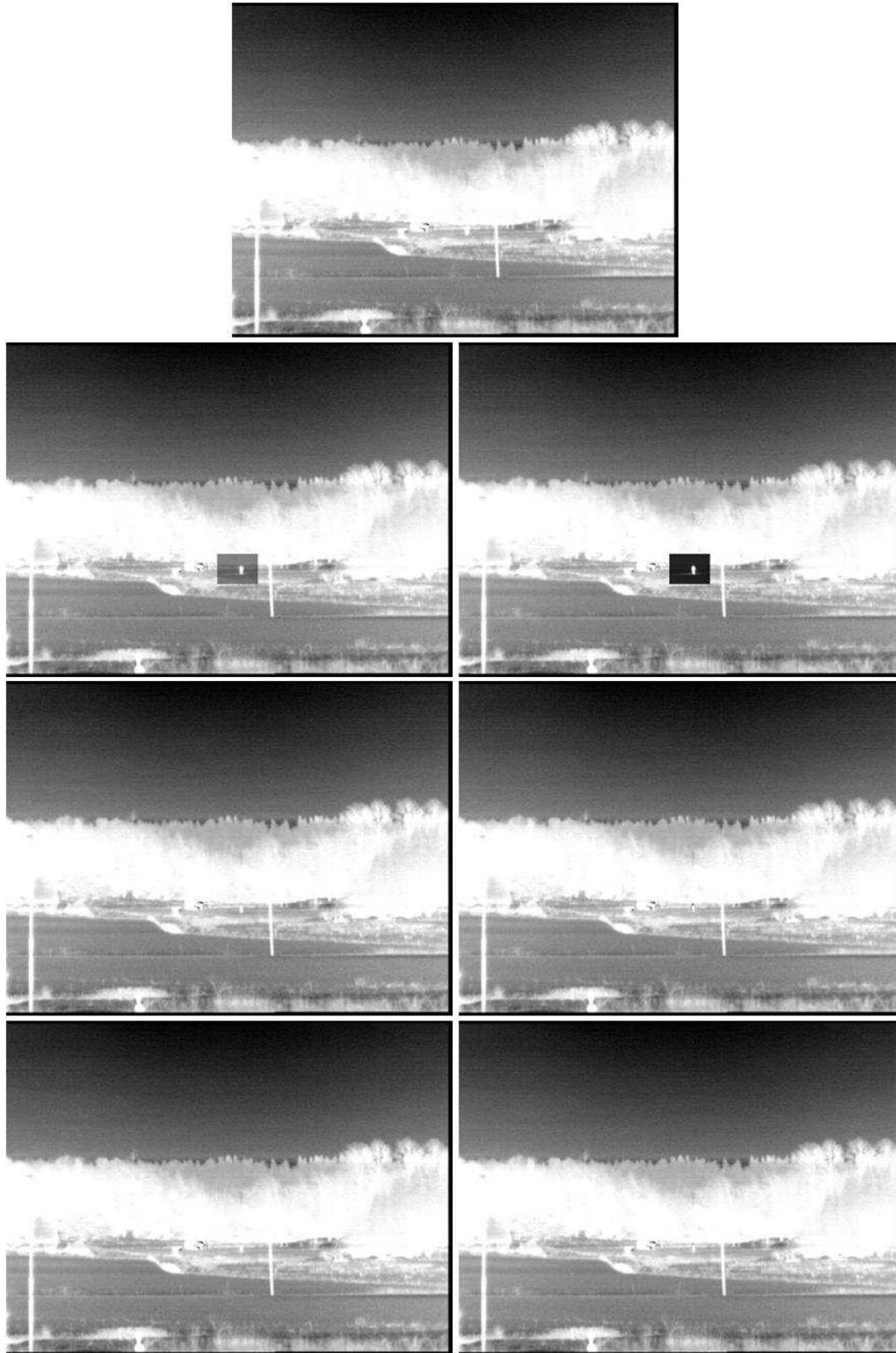


Figure 10: examples of final fusion results F top to bottom left to right, IR (as provided by FOI), 1 (PC5), 3 (PWP), 4 (PCM), 5 (RPF), 8 (RWM), 7 (RC5). The numbers and acronyms are corresponding to the human conspicuity test in section 6.1 and the results in Table 1.

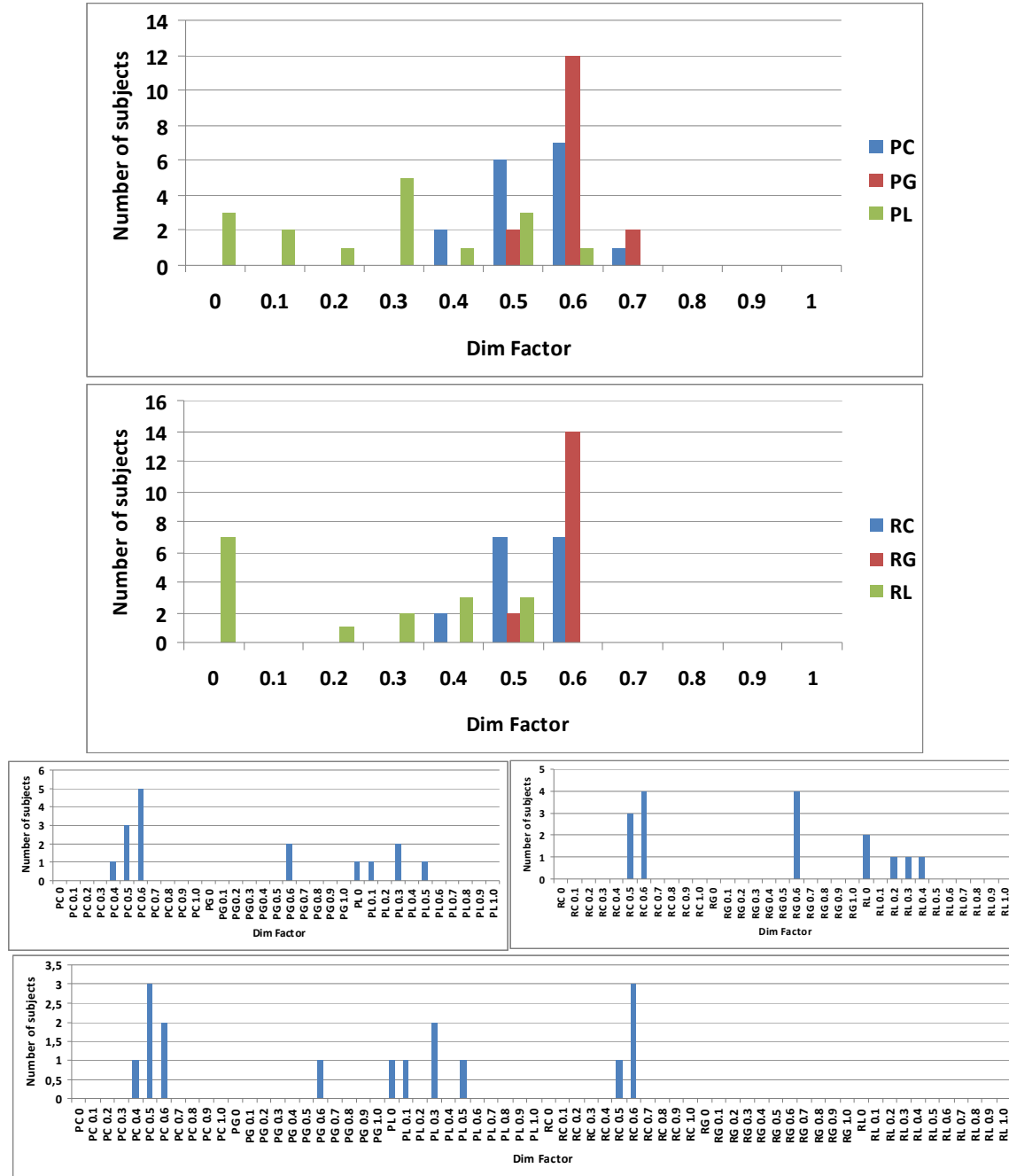


Figure 11: results of the human experiment for dimming on scenario 7 for both the pixel based and region based approach and for all three dimming methods. PC is pixel based with contextual dimming, PG is pixel based with global dimming and PL is pixel based with local dimming. RC is region based with contextual dimming, RG is region based with global dimming and RL is region based with local dimming. The top chart is distribution of optimal dim factor for all three dimming methods on the pixel based approach, defined in step 9.1. The second chart from the top is distribution of optimal dim factor for all three dimming methods on the region based approach, defined in step 9.1 as well. The third chart from the top on the left is distribution of optimal dimming for the pixel based approach, defined in step 9.2. The third chart from the top on the right is distribution of optimal dimming for the region based approach, defined in step 9.2 as well. The chart on the bottom shows the distribution of the overall preferred dimming, defined in step 9.3.

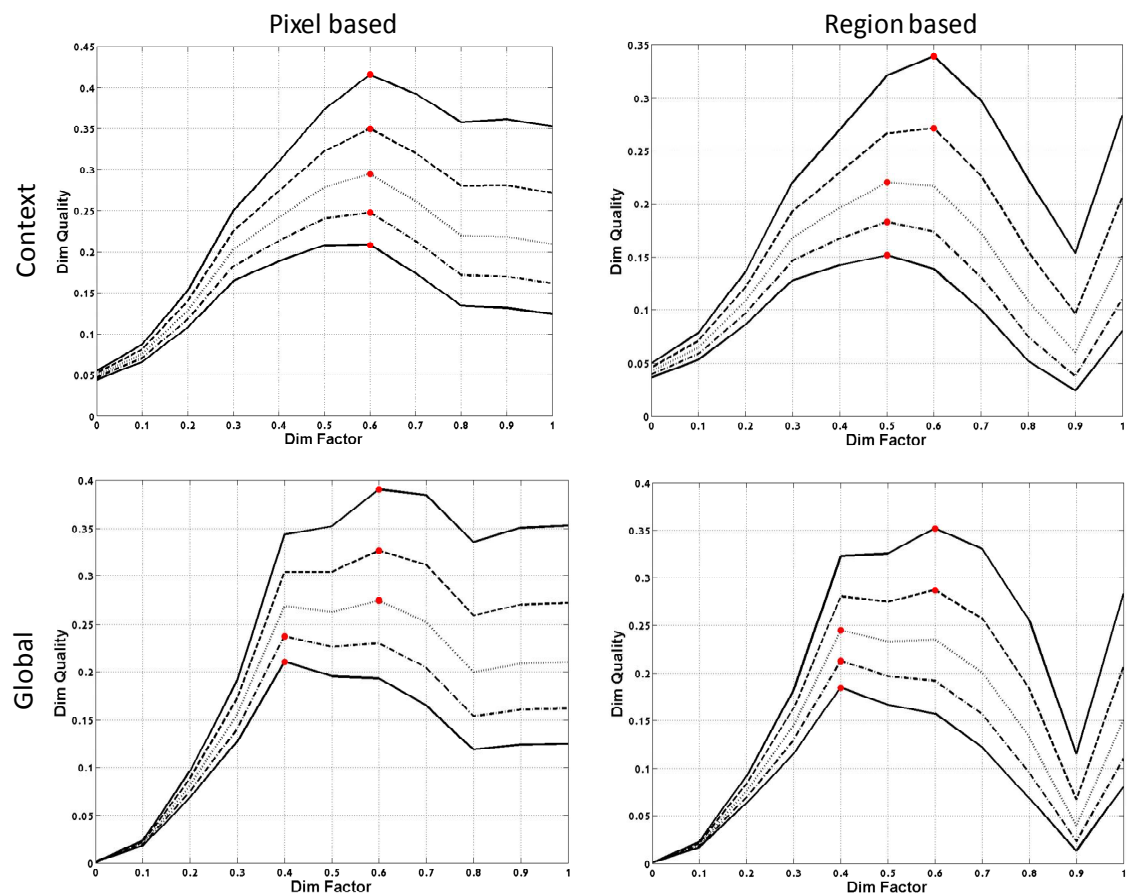


Figure 12: dim quality for different steering parameter settings (top to bottom 0.4 to 0.8 in steps of 0.1) against dim factor for global and contextual dimming on both the pixel based as well as the region based approach. Optima are indicated with red dots.

Part 3: Supplementary material

1 Introduction

This document provides supplementary material for the article *Semi-hidden target recognition in gated viewer images fused with traditional thermal IR images*. It provides additional explanation on several topics in order to better understand the content of the article, which is required for the reader who is less familiar with Image Processing and Electro-Optical theory. Topics are provided in the order it appears in the article. The topics that will be discussed are the Laser Range Gated Viewer principle, image fusion using wavelet decomposition and the image decomposition modules that are used in the study.

2 Laser Range Gated Viewer

Figure 13 shows the principle of a Laser Range Gated Viewer (GV). This laser system consists of a pulsed laser in the reflective infrared (IR) band and a receiver with detector of a resolution similar to a normal IR camera. The laser pulse which is transmitted by the laser will be reflected by all objects/material in the path it propagates. Normally all the reflected laser energy will be received and the time of flight says at what distance the objects are located as the laser pulse travels with the speed of light. This is how a Laser Range Finder works. The Laser Range Gated Viewer is in two ways adapted. First of all, it has detector with a high resolution, which is combined with optics that provide a certain field of view. This makes it possible to create an image of a scene like a traditional camera in which the pixel intensity represents the received (reflected) laser energy. Secondly, the receiver is able to look at reflected laser energy at a specific time and thus by making use of the time of flight the receiver it is able to receive reflected energy from a specific distance. All other reflected energy will be neglected. When the location of the target is known, it is possible to lay a *gate* in time-distance over the target. The length of the gate in time (or the time the receiver is “open”) defines the deepness of the gate in distance and the time the gate starts (or the receiver “opens”) defines the distance the system is looking at. In this way only information that lays within the range gate, and thus from the range of the object, will be received for which the system provides an image. All reflected energy from objects and material in between the system and target and behind the target will be neglected whereas reflected energy from the target will be available in the image. It is obvious that if the target is behind an object the laser pulse will not reach the target, however, all visible parts of a target will provide reflected energy. Therefore, this technique is especially interesting in scenarios with targets merged in the scene, e.g. in between trees (semi-hidden targets).

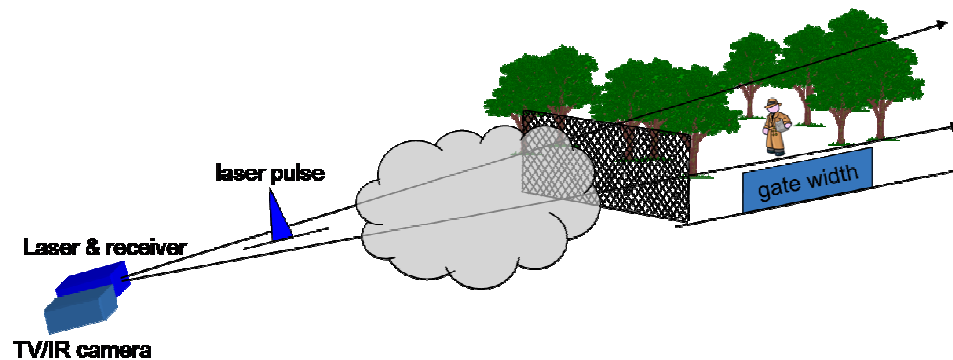


Figure 13: principle of the Laser Range Gated Viewer combined with a traditional TV/IR camera. The camera will provide an image that contains all scene information whereas the Laser Range Gated Viewer provides an image that only contains information from the range gate. The drawing is provided by TNO.

3 Image fusion using Wavelet decomposition

Figure 14 shows the fusion scheme for an algorithm that uses wavelet transform decomposition. The idea is similar when using all different multi-resolution decomposition methods.

The transform is used to decompose the initial images in high frequency components and a low frequency component. The wavelets transform filters the image into a low frequency component that approximates the image and three directional high frequency components that contain details: horizontal, vertical and diagonal. Next, the low frequency component is down sampled and after being filtered again a new scale is created with the same combination of components of a smaller size. This is repeated for the amount of requested scales. The example in Figure 14 contains three scales: the low frequency component is in the left upper corner of the decompositions and the other 9 components represent the high frequencies in the three different scales. Fusion is applied on the separate components. The low frequency component is fused with its counter part of the other image and each high frequency component is also fused with its counterpart. Fusion can be applied using the same or separate fusion rules (mathematical operations) for the low and high frequency components, as long as the fusion rules for the high frequency components are the same for each pair of components. Fusion rules can be an average weighting of the corresponding components, a selection of the entire component of either one of the images or a selection at each location in the component. The article describes several fusion rules. After fusion one image decomposition is created. The fused image is obtained by applying the inverse of the wavelet transform.

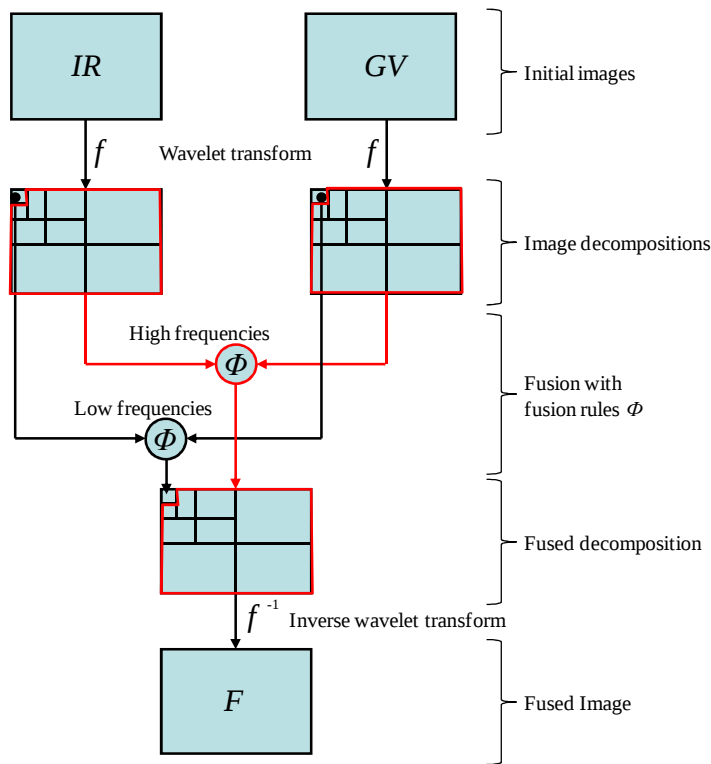


Figure 14: Image fusion using the wavelet image decomposition, reproduced from [1]. IR and GV are the input images, f is the transform with its inverse transform f^{-1} and Φ is the fusion rule.

4 Image decomposition modules

Wavelet module

The wavelet transformation has been known as a good decomposition method for fusion in which one is capable to fuse both directional high- and approximation low frequencies by distinctive fusion rules in order to preserve details and background information. Basic theory for wavelet theory is provided by S.G. Mallat [2]. A disadvantage of the basic wavelet transform is the limited amount of directional information; only horizontal, vertical and diagonal high frequency components. Another disadvantage is that the digital wavelet transform is not shift invariant due to sub sampling. Both issues are addressed in the Dual Tree Complex Wavelet Transform (DT-CWT) developed by N. Kingsbury [3], [4], [5]. Examples for the use in image fusion are provided by Nikolov et al. [1] and Lewis et al. [6]. In the DT-CWT two decimated trees are produced, one containing the odd samples and one the even samples after the first level filtering. In each tree in the next levels the samples are taken even and odd alternately. The impulse response of the bi-orthogonal filters represents a real and imaginary part of a complex wavelet. A result of the DT-CWT is six directions for the high frequency coefficients rather than three: $\pm 15^\circ$, $\pm 45^\circ$ and $\pm 75^\circ$.

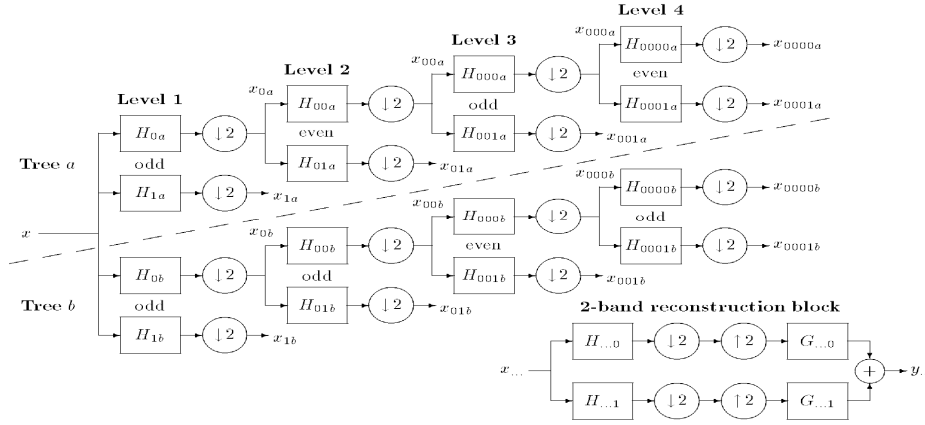


Figure 1: Dual tree of filters for the complex wavelet transform.

Figure 15: Complex wavelet decomposition scheme with dual tree and alternated odd/even samples, taken from [4]

Curvelet module

The wavelet transformation still has some disadvantages with respect to curving edges. The wavelet is especially suitable for representing object edges with point singularities, but has trouble representing singularities along smooth curves (higher dimensional structures). The curvelet and fast discrete curvelet proposed by Candès et al. [7] [8] is able to sparsely represent objects which display smoothness except for discontinuity along a general curve with bounded curvature. It is a multi scale pyramid with significantly more directions and positions at each length scale and needle-shaped elements at fine scales. The translation and rotation of a mother curvelet represent the high frequency coefficients and in the decomposition they are placed around a central low frequency coefficient. In the curvelet transform the amount of directional information is even more increased compared to the DT-CWT and is therefore a suitable method for the goal of the study as well. Deng et al [9] showed an example of image fusion using the curvelet transform, in which they show it is also suitable for noisy images.

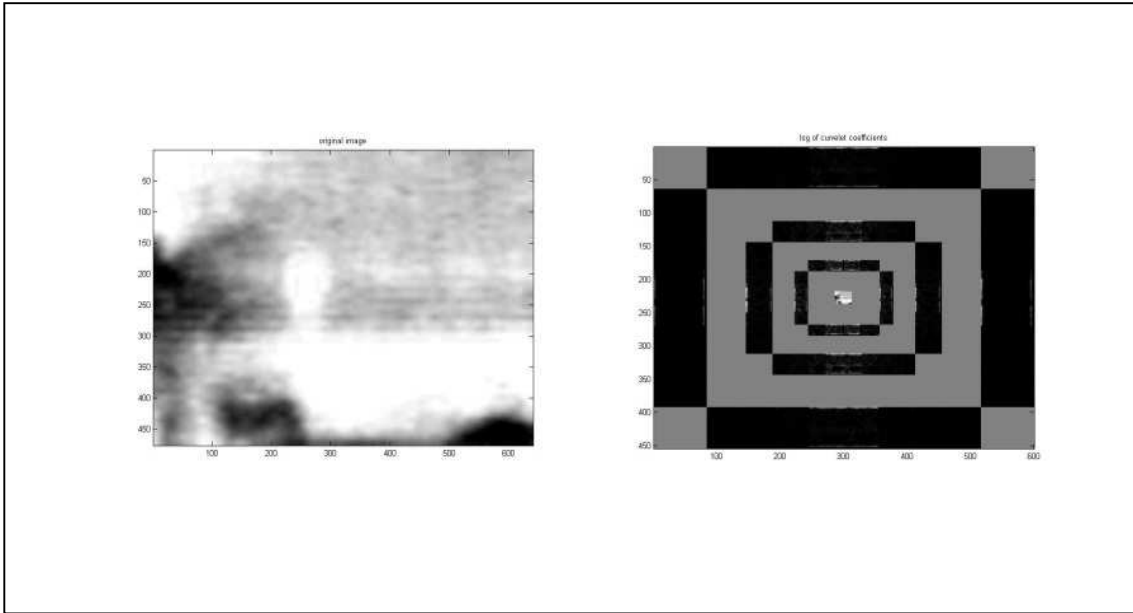


Figure 16: an example image decomposition using the curvelet, on the left the initial image and on the right the decomposition, created using the curvelet transform of [7] and [8]. In the center is the low frequency component located and around it the high frequency components are placed.

Contourlet module

Another interesting multi resolution transform representing more directional information is the contourlet proposed by Do and Vetterli [10], which moreover has flexibility in choosing the number of directions at each level. As it immediately constructs its decomposition in a discrete domain it is also less complex. An example for use in region based image fusion is shown by Ibrahim and Wirth [11] for fusion visible and IR images. The contourlet makes use of a double filter bank: a Laplacian pyramid (LP) capturing the low frequencies and a directional filter bank (DFB) that represents the directional high frequency components. The DFB consist of a two channel quincunx filter bank dividing the spectrum into horizontal and vertical directions and a shearing operator that is used for rotation of the image. In the decomposition again the low frequency component is the central part and the high frequency component is divided in directional sub bands located around the low frequency component. Like the curvelet this transform is suitable for representing discontinuities along smooth curves at object edges or contours.

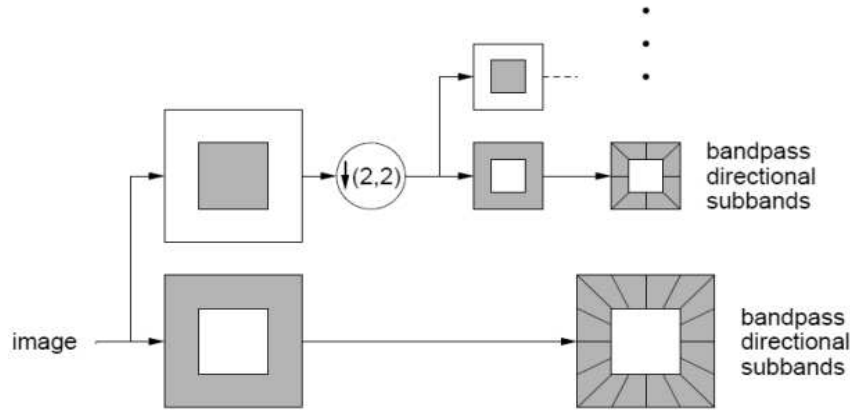


Figure 17: decomposition scheme of the contourlet (figure 7 from [10])

5 References

- [1] S.G. Nikolov, P.R. Hill, D.R. Bull, C.N. Canagarajah, "Wavelets for image fusion," in *Wavelets in Signal and Image Analysis, Computational Imaging and Vision Series*. Dordrecht, the Netherlands: Kluwer Academic Publishers, 2001, pp. 213-244.
- [2] S.G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693, 1989.
- [3] N.G. Kingsbury, "The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancement," in *Proceedings of EUSIPCO 98*, Rhodes, 1998, pp. 319-322.
- [4] N.G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Journal of Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234-253, May 2001.
- [5] I.W. Selesnick, R.G. Baraniuk, N.G. Kingsbury, "The Dual-Tree Complex Wavelet Transform: A coherent framework for multiscale signal and image processing," *IEEE Signal Processing Magazine*, pp. 123-151, November 2005.
- [6] J.J. Lewis, R.J. O'Callaghan, S.G. Nikolov, D.R. Bull, "Pixel- and region-based image fusion with complex wavelets," *Elsevier Information Fusion*, vol. 8, pp. 119-130, 2007.
- [7] E.J. Candès, D.L. Donoho, "Curvelets - A Surprisingly Effective Nonadaptive Representation For Objects with Edges," in *Curves and Surfaces*. Nashville, Tennessee: Vanderbilt University Press, 2000, pp. 105-120.
- [8] E.J. Candès, L. Demanet, D.L. Donoho, L. Ying, "Fast discrete curvelet transform," *Multiscale Modeling and Simulation*, vol. 5, no. 3, pp. 861-899, 2006.
- [9] C. Deng, H. Cao, C. Caob, S. Wang, "Multisensor image fusion using fast discrete curvelet

transform," in *Proceedings of SPIE*, vol. 6790 679004-1, 2007.

[10] M.N. Do, M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091-2106, 2005.

[11] S. Ibrahim, M. Wirth, "Visible and IR data fusion technique using the contourlet transform," in *IEEE International Conference on Computational Science and Engineering*, vol. 2, 2009, pp. 42-47.

Part 4: Report / research log

1 Introduction

This report / research log is meant to provide an overview of all the work performed for this study, by means of a log that also contains results including results that are not shown in the article, for instance the registration solution. It also contains a list of the matlab code created in this study.

Although the first 4 sections are written in English, the research log is written in Dutch.

First the goal of the study is provided, followed with the research questions. Section 4 provides the study approach. Section 5 discusses the planning and section 6 provides the log of the study including important results. Results that are captured with large tables or images as well as all the matlab codes are provided in appendices.

2 Study goal

The goal of this study is to find out if semi-hidden target recognition will be improved when adding Laser Range Gated Viewer (GV) in a multi sensor set-up and to define a fusion method for GV images with IR images that improves the recognition task of humans.

3 Research questions

3. Does a GV improve semi-hidden target recognition by humans in a multi sensor GV+IR set-up?
4. What is an optimal fusion method for GV with IR images in order to improve semi-hidden target recognition by humans?

4 Study approach

The IR and GV images will be registered using existing techniques that are adapted to the scenario. Based on literature fusion algorithms will be designed. An Image Quality Metric (IQM) will be used to select the optimal fusion algorithm from those considered. This IQM will be improved with better saliency metrics. The IQM will be validated using human conspicuity experiments. Next, a background dimming algorithm is designed. The preferred amount of dimming is also defined using a human experiment. This experiment will also be used to validate a designed IQM for dimming, or Dim Quality Metric (DQM). In a final experiment the influence of moving objects on the preferred amount of dimming will be investigated. Based on the results an optimal approach for multi-sensor semi-hidden target recognition using GV and IR images can be described and an answer to the question if GV improves the recognition task of semi-hidden targets can be given as well.

5 Verantwoording activiteiten

Week:	6	Datum:	07-02-2011
Activiteiten:	- Opstellen logboek - Opstellen discussie document: studiedoelstelling, outline en onderzoeksvragen		
Doel:	- Discussie studie aanpak TNO		
Opmerkingen:	-		
Resultaten:	-		

Week:	6	Datum:	11-02-2011
Activiteiten:	- Afronden literatuuronderzoek - Bezoek TNO: bespreken werkplek		
Doel:	- Regelen werkplek TNO		
Opmerkingen:	-		
Resultaten:			

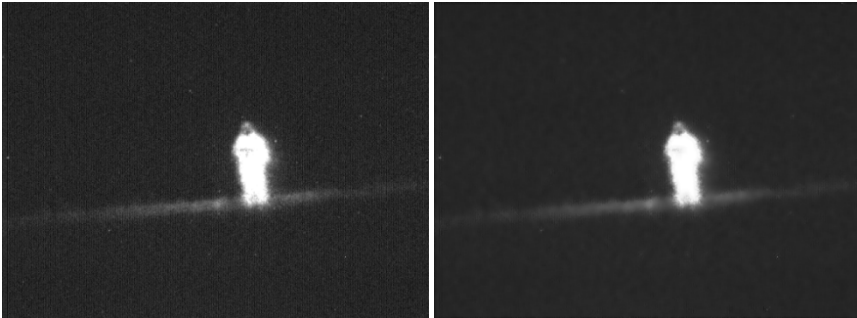
Week:	7	Datum:	14-02-2011
Activiteiten:	- Afronden doelstelling, onderzoeksvragen en outline en aanleveren TNO		
Doel:	- Discussie TNO		
Opmerkingen:	-		
Resultaten:	- Nieuw outline document		

Week:	7	Datum:	20-02-2011
Activiteiten:	- Aanpassen doelstelling, onderzoeksvragen en outline n.a.v. discussie TNO - Start plan van aanpak: nieuw flow chart		
Doel:	- Opstellen doelstelling, onderzoeksvragen, study outline - Start Plan van aanpak		
Opmerkingen:	- Na bestuderen data dient opnieuw gekeken te worden of doelstelling relevant is en of onderzoeksvragen beantwoord kunnen worden.		
Resultaten:			

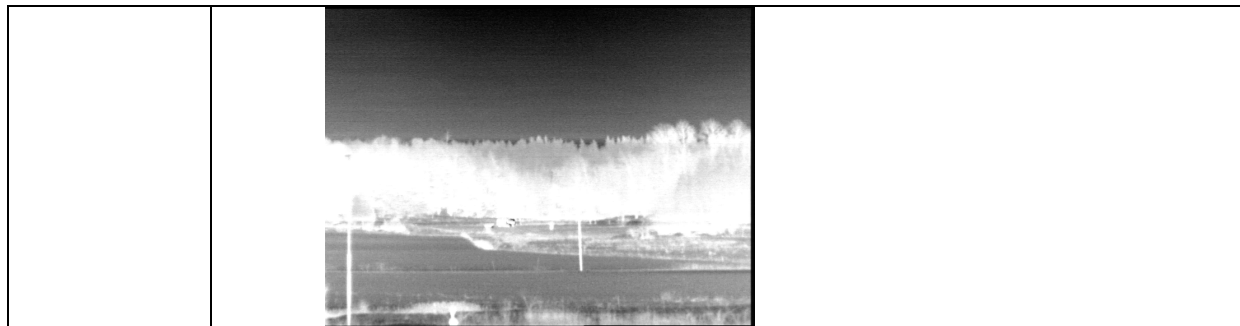
Week:	8	Datum:	27-02-2011
Activiteiten:	- Download en installatie Matlab - Download en installatie Visual Studio C++ - Download elastix registratie software - Bekijken IR images in matlab - Opzet plan van aanpak		
Doel:	- Voorbereiden thuiswerk - Bestuderen data		
Opmerkingen:	- IR image sequences bestaan uit >>50 beelden in tijd en kunnen beeld voor beeld bekeken worden		
Resultaten:	- Opzet plan van aanpak - Inzicht data		

Week:	9	Datum:	01-03-2011
Activiteiten:	- Presentatie voortgangsmeting voorbereiden		
Doel:	- Voorbereiden voortgangsmeting		
Opmerkingen:	-		
Resultaten:	-		

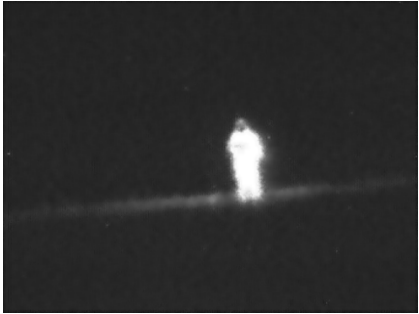
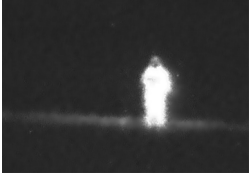
Week:	9	Datum:	02-03-2011
Activiteiten:	- Bestuderen data		
Doel:	- T.b.v. dataselectie en vragenlijst		
Opmerkingen:	-		
Resultaten:	-		

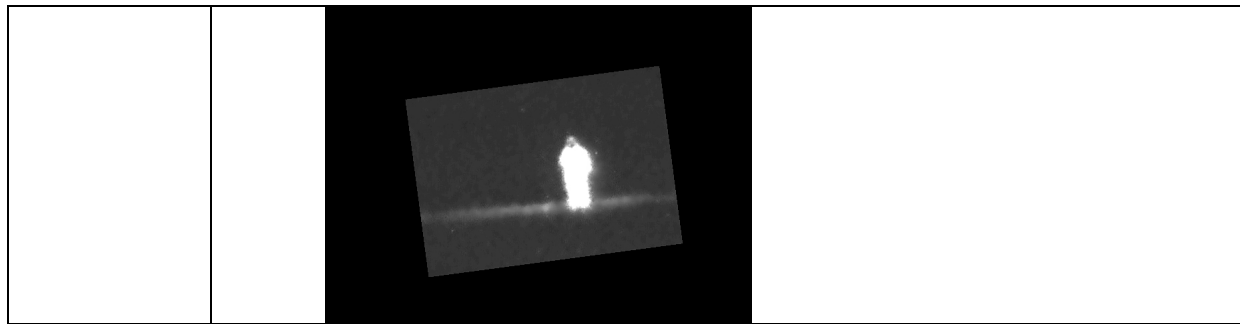
Week:	9	Datum:	06-03-2011
Activiteiten:	- Opstellen vragenlijst over data (t.b.v. FOI) - Image pre-processing GV images: • Matlab imadjust t.b.v. contrast enhancement • Matlab wiener filter size 10 t.b.v. noise removal - Start pre-processing IR images		
Doel:	- T.b.v. dataselectie - Benodigd ter voorbereiding registratie		
Opmerkingen:	- IR nog geen dataset voorbereid		
Resultaten:	- Alle GV images: contrast enhancement (links) en simpele noise removal (rechts). Alles TIFF formats. 		

Week:	10	Datum:	07-03-2011
Activiteiten:	- IR dataselectie - IR pre-processing: alleen imadjust indien noodzakelijk		
Doel:	- T.b.v. dataselectie - Benodigd ter voorbereiding registratie		
Opmerkingen:	- Gebruik gemaakt van m-file bij data: readRawImage		
Resultaten:	- Alle IR frames als tiff beschikbaar		



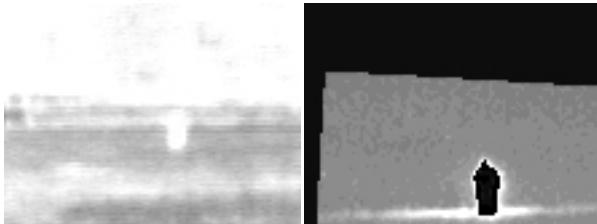
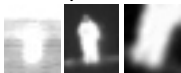
Week:	10	Datum:	08-03-2011
Activiteiten:	- IR videostream gemaakt - Installeren elastix (in c++) en bestuderen handleiding		
Doel:	- T.b.v. dataselectie: bestuderen videostream en uitlijnen in tijd - Elastix t.b.v. registratie		
Opmerkingen:	- Object t.b.v. classificatie beweegt niet in IR stream: dus uitlijnen in tijd niet noodzakelijk		
Resultaten:	- Video stream in avi en wmv - Werkende elastix		

Week:	10	Datum:	13-03-2011
Activiteiten:	- Testen en leren elastix registratie software: GV beeld als fixed image en een geroteerd beeld met andere scaling als moving image		
Doel:	- T.b.v. registratie: aanleren gebruik elastix en testen werking elastix		
Opmerkingen:	- Test is single modality. Diverse testen uitgevoerd om aan te leren en vast te stellen of gebruik elastix zinvol is. Beste resultaat voor single modality GV weergegeven.		
Resultaten:	- Elastix met single modality functioneert naar behoren - Links: testfixed image, rechts test moving image (rotatie + scaling)   Testresult elastix met MI en rigid transformation:		

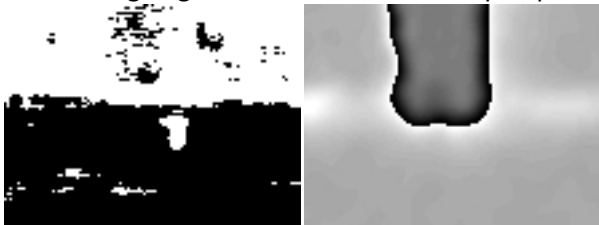


Week:	11	Datum:	17-03-2011
Activiteiten:	- Student meeting - Bestuderen antwoorden vragenlijst Zweden		
Doel:	- T.b.v. registratie		
Opmerkingen:	- Zweedse collega's hebben geen extra data van voertuigen - FOV GV = 3.8 mrad - FOV IR = 2.4° - Update rate GV = 10Hz - IR en GV niet in tijd uitgelijnd - Hoek optische assen niet bekend: dus registratieproces noodzakelijk		
Resultaten:			
Week:	11	Datum:	18-03-2011
Activiteiten:	- Uitwerken plan van aanpak registratie - Uitvoeren test 1 en 2 (zie lijst testen in vak resultaten)		
Doel:	- T.b.v. registratie		
Opmerkingen:	- Op basis van gedane literatuurstudie - En op basis mogelijkheden elastix - Testen in volgorde stapsgewijs doorlopen totdat succesvolle methode gevonden is. - In principe: MI beste voor multi-modal en er is sprake van translation, rotation en scaling (dus rigid of affine meest voor de hand liggend) - Volgens literatuurstudie: • mod 2 is small area of mod 1 best method is Normalized Gradient Fields with "cartoon image" method (table 1 en ref 26 en 28 literatuur studie) • mod 2 some structure and mod 1 rich of structures: Non-rigid transformation using MI and CC with "pseudo modality", MI or Normalized Gradient Fields (see table 1, ref 29, 26 of literatuurstudie)		
Resultaten:	Plan van aanpak: IR fixed image, GV moving image. Steeds diverse transformatie methoden proberen: rigid, affine en non-rigid. Tevens parameters aanpassen. - Test 1: MI en NMI op initiele IR en GV beelden - Test 2: MI en NMI op "patch" IR (region van GV) en GV beeld - Test 3: MI en NMI op "patch" IR en downsampled GV (gelijke resolutie als IR patch)		

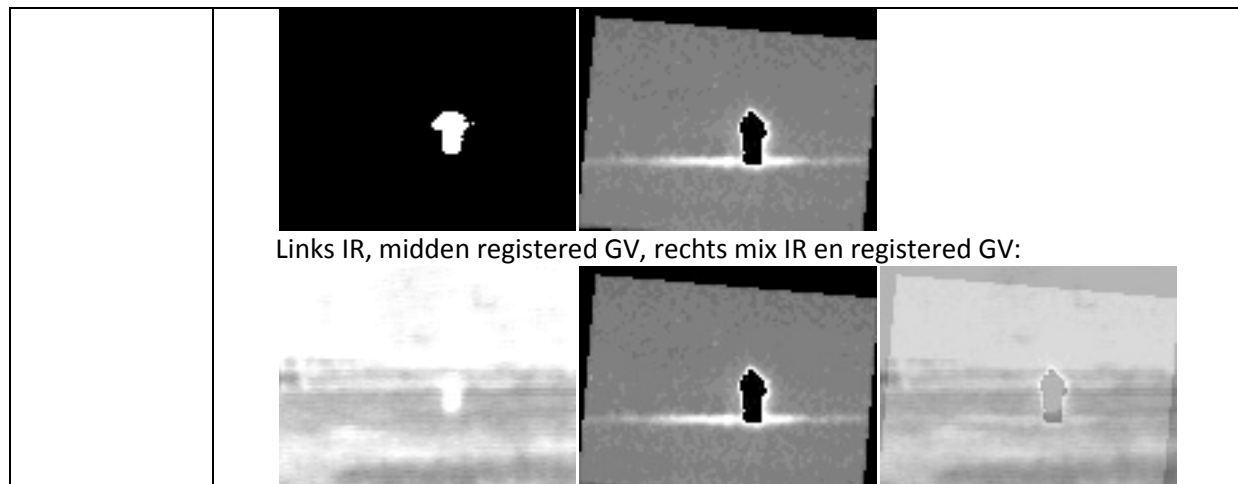
	- Test 4: MI op patch IR (alleen persoon) en downsampled patch GV (ook alleen persoon) - Test 5: MI op initiele patch IR en down sampled GV met erodemask (avoid bad karma from artificial edge) - Test 6: gebruik maken van corresponding points (dit heeft niet de voorkeur omdat op iedere image corresponding points aangegeven moeten worden) - Test 7: gebruik maken van binary mask image op IR - Test 8: gebruik maken van image processing techniques (gradient, derivative/2nd order derivative, Laplacian etc en na filtering registratie uitvoeren) - Test 9: segmentation van object en Kappa statistics als similarity metric (specifiek voor binary images) - Test 10: indien geen van eerdere testen werken zoeken naar registratie methoden ref 26, 28 en 29 van literatuurstudie Zodra meest succesvolle methode gevonden is: toepassen op alle data en indien gebruik gemaakt van down sampled versie van GV, met transformix tool toepassen op initiele resolutie GV. Indien patch gebruikt is binnen GV beeld is met dit laatste enige voorzichtigheid nodig i.v.m. rotatiepunt. Dat moet dan wel gelijk zijn qua positie in zowel GV patch als gehele GV plaatje. Uitvoeren test 1 en 2: - Test 1: NOK (NOK = not OK) - Test2: NOK
--	---

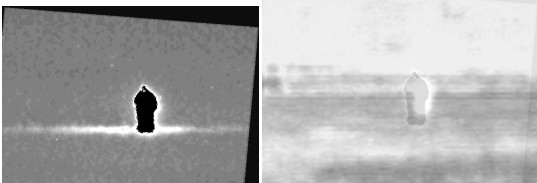
Week:	11	Datum:	20-03-2011
Activiteiten:	- Uitoeren registratie testen: test 3 t/m 5		
Doel:	- T.b.v. registratie		
Opmerkingen:	- Resultaten lijken nog niet op enige goede richting van registratie - Intensiteit lijkt ook aangepast te worden door elastix...???		
Resultaten:	- Test 3: NOK - Test 4: NOK - Test 5: NOK "Beste" resultaat tot nu toe test 3 (left patch IR, right result, the GV image is cf table 06-03-2011):  "Beste" resultaat tot nu toe test 4 (left IR patch of patch, middle GV patch, right result): 		

Week:	12	Datum:	22-03-2011
Activiteiten:	- Uitoeren registratie testen: test 6		
Doel:	- T.b.v. registratie		
Opmerkingen:	- Inlezen van points cf handleiding gaat steeds fout. Krijg input txt file niet in juiste format zodat elastix deze meeneemt. Aangezien ik deze methode uiteindelijk niet ga gebruiken laat ik deze test voor wat het is.		
Resultaten:	- Test 6: NOK		

Week:	12	Datum:	23-03-2011
Activiteiten:	- Uitoeren registratie testen: test 7 (deel 1) - Voorbereiden registratie test 8: installeren DIPimage toolbox en testen diverse filters op beelden.		
Doel:	- T.b.v. registratie		
Opmerkingen:	- Het lijkt er op dat mask kan gaan werken. In vervolg test dient mask specifiek waarde 1 alleen voor object te hebben. - Het lijkt erop dat t.b.v. test 8 filters alleen in x-richting gaan werken. In y richting leveren gradient/derivative te veel informatie in IR t.o.v. GV.		
Resultaten:	- Test 7: tot nu toe NOK, maar mask werkt wel in elastix (zie opmerkingen en resultaat hieronder). Aan resultaat is te zien dat het kleinste verschil ontstaat als het object uit de GV vergroot wordt en over de bomen in IR wordt gelegd. Resultaat voor mask op IR patch(links mask, rechts resultaat): 		

Week:	13	Datum:	28-03-2011
Activiteiten:	- Bijwerken logboek - Uitoeren registratie testen: test 7 (deel 2) - Voorbereiden registratie test 8: testen diverse filters op beelden.		
Doel:	- T.b.v. registratie		
Opmerkingen:	- Mask is momenteel nog m.b.v. photoshop gecreeerd op threshold in DIPimage. Aangepaste Matlab algoritme moet nog gemaakt worden. - Resolutie lijkt door elasticx ook aangepast te worden. Reden nog niet gevonden. - Correcte registratie nog neit op initiele resolutie GV image gelukt. Vervolg stap - Wel verder gegaan met juiste filter methode voor test 8 te vinden, maar gezien resultaat met mask ligt prio bij test 7. Zodra test 7 succesvol is op initiele resolutie: uitvoeren registratie gehele dataset.		
Resultaten:	- Test 7: lijkt OK, maar allen op downsampled GV beeld. Overgebleven issues zijn verandering in intensiteit en uitvoeren registratie op initiele resolutie GV beeld. Links mask, rechts result:		



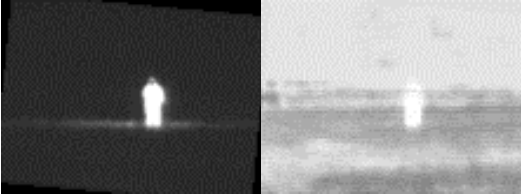
Week:	13	Datum:	03-04-2011
Activiteiten:	- Vervolg uitvoeren registratie test 7 (deel 3)		
Doel:	- T.b.v. registratie		
Opmerkingen:	- Test 7a: uitvoeren transformix op initiele resolutie van GV beeld - Test 7b: uitvoeren elastix op initiele resolutie GV beeld		
Resultaten:	- Test 7a: NOK. Lijkt in eerste instantie te kloppen (uitvoeren juiste transformaties), maar resultaat is niet precies te leggen op IR beeld Links resultaat transformix, rechts bovenop IR  - Test 7b: NOK (volledig verkeerde registratie, dus downsamen van GV beeld noodzakelijk)		

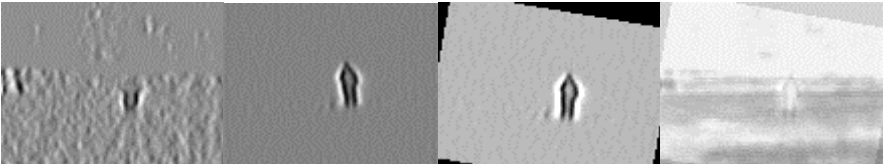
Week:	14	Datum:	04-04-2011 @TNO
Activiteiten:	- Afronden administratieve zaken voor werkplek TNO - Opstarten computer account TNO - Maken presentatie voortgang		
Doel:	- Regelen werkplek TNO - T.b.v. voortgangsmeting		
Opmerkingen:			
Resultaten:	- N.v.t.		

Week:	14	Datum:	07-04-2011
Activiteiten:	- Voortgangsmeting		
Doel:	- Bewaken voortgang		
Opmerkingen:	- Up-samplen IR (bi-linear) i.p.v. GV down-samplen - Probeer rand detectie voor registratie (gradient) - Intensiteit probleem oplossen (eventueel vragen Marius Staring LUMC)		
Resultaten:	- N.v.t.		

Week:	15 t/m 18	Datum:	Vakantie
Activiteiten:	- Ideeën voor fusie algoritme op papier uitwerken op basis literatuur		
Doel:	- T.b.v. fusiealgoritme		
Opmerkingen:	-		
Resultaten:	Eerste ideeën: - beeldfusie met behoud IR achtergrond en “highlighten” van human/object. Herkenning door mens is optimaal bij versterken directional info en behoud achtergrond, maar object moet duidelijk naar voren komen. In IR “zinkt” object weg in de achtergrond daarom is een region based methode (regio 1 is object, regio 2 is achtergrond/omgeving) met priotisering regio 1 GV en regio 2 IR. Diverse fusie algoritmes toepassen op regio van object: prio GV, weighted, contourlet/curvelet/DT-CWT, ICA. Vergelijken met basis fusie algoritmes op gehele beeld. IR achtergrond geen GV in fuseren maar prio IR en om object te highlighten achtergrond “dimmen”. IQM toepassen op gehele beeld én gebied rondom object, zgn patch 1 en patch 2 (p1 en p2). IQM is optimaal voor gehele beeld als er niet gedimmed wordt, maar object verdwijnt dan in de achtergrond. IQM is optimaal in gebied rondom object bij maximaal dimmen van achtergrond, maar dan verdwijnt de achtergrond informatie. Dus zoeken naar optimum met herkenning object en behoud achtergrond. Mogelijk $\max\{IQM_object * IQM_achtergrond\}$.		

Week:	19	Datum:	09-05-2011 @TNO
Activiteiten:	- Vervolg uitwerken fusie theorie - Vragen uitzetten bij M Staring LUMC t.b.v. registratie - Vragen uitzetten bij A Toet TNO t.b.v. gebruikte basis fusie algoritmes - Internet zoektocht fusie toolboxes in matlab		
Doel:	- T.b.v. oplossen registratie problemen en start fusie algoritme		
Opmerkingen:	- Verdeling werk: fusie@TNO en registratie@home - Basis toolbox: Matifus - ICA: DTU toolbox - Curvelet: candes et al. Curvelet.org - Contourlet: minh n. do contourlet toolbox - DT-CWT: website Nick Kingsbury		
Resultaten:			

Week:	19	Datum:	10-05-2011
Activiteiten:	- Oplossen intensiteitsprobleem registratie a.d.h.v. antwoord M Staring		
Doel:	- T.b.v. registratie		
Opmerkingen:	- Test 7c: is herhaling test 7 maar dan met 8bit/pixel voor GV i.p.v. 16bit/pixel		
Resultaten:	- Test 7c: OK v.w.b. oplossen intensiteits probleem. NOK: positie en rotatie OK, maar scaling klopt nog niet. Links resultaat, rechts bovenop IR 		

Week:	19	Datum:	15-05-2011
Activiteiten:	- Vervolg registratie test7 met andere mask - Uitvoeren registratietest 8 met randdetectie		
Doel:	- T.b.v. oplossen registratie problemen		
Opmerkingen:	- Test 7: rechthoekige maskeringen rondom object gebruikt - Test 8: randdetectie lastig in IR. Uiteindelijk na stretchen grijswaarden en gebruik second derivative met nogmaals stretching grijswaarden een bruikbare edge representatie verkregen. Second derivative alleen in x-richting om ongewenste horizontale lijnen van overgang bomenrand en grond te vermijden.		
Resultaten:	- Resultaat test 7d: NOK, gelijk aan 7c dus alleen scaling nog niet goed. Links mask, midden resultaat, rechts bovenop IR  - Resultaat test 8: NOK. Gelijk aan resultaat test 7c en d. Scaling niet goed, iets te veel rotatie en tevens lukt uitvoeren transformix niet op initiele GV beeld. Links edgerepresentatie IR, middenlinks edgerepresentie GV, middenrechts resultaat en rechts bovenop IR  Vervolg registratie: - Oplossen scaling door gebruik te maken van a-priori FOV en aanname CoM voor object in IR en GV gelijk en vervolgens toepassen rigid transformaties (dus geen scaling). - Ik zie geen verbetering door gebruik te maken van edge representaties dus ik blijf bij maskeren.		

Week:	20	Datum:	16-05-2011 @TNO
Activiteiten:	- Bijwerken logboek - Vervolg: bestaande fusie toolboxes zoeken (internet) - Zoeken papers IQM - Bestaande matlab algoritmes selecteren voor: segmentatie en registratie		
Doel:	- T.b.v. oplossen registratie problemen, start fusie algoritme en vaststellen performance methodieken		
Opmerkingen:	- Toolboxes gevonden maar nog niet kunnen downloaden		
Resultaten:	- Papers IQM: Gemma Piella, Xydeas & Petrovic en IQI (Wang & Bovik) - Lijst bestaande matlab algoritmes : eigenschappen van regions als centre of mass en bounding box (regionprops), regionlabels (bwlabel). Allen t.b.v. region segmentation voor fusie en automatisch bepalen van mask voor registration.		

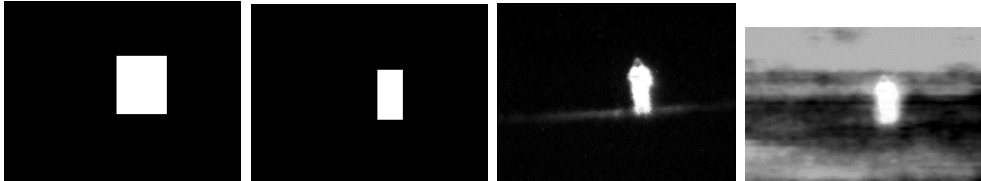
Week:	21	Datum:	29-05-2011 Thuis
Activiteiten:	- Op papier uitwerken auto algoritme preparatie beeldregistratie		
Doel:	- T.b.v. oplossen registratie problemen		
Opmerkingen:	-		
Resultaten:	-		

Week:	22	Datum:	30-05-2011 @TNO
Activiteiten:	- Download sessie fusie toolboxes		
Doel:	- T.b.v. beeldfusie algoritmes		
Opmerkingen:	- Alleen DT-CWT toolbox (prof Kingsbury) nog niet beschikbaar. Aanvraag via e-mail		
Resultaten:	Beschikbare toolboxes voor matlab: - MATIFUS wavelets & Laplacian Pyramids: CWI - ICA toolbox: MIALAB - Curvelab : curvelet.org, Candes et al - Contourlet toolbox : Minh-Do - Gabor wavelets algoritme reeds beschikbaar in matlab zelf		

Week:	22	Datum:	31-05-2011 Thuis
Activiteiten:	- Bestuderen papers 7, 8 en 9		
Doel:	- T.b.v. IQM voor evaluatie recognition by humans		
Opmerkingen:	-		
Resultaten:	-		

Week:	22	Datum:	03-06-2011 Thuis
Activiteiten:	- Deel 1 van registration algoritme maken: creation of binary masks		
Doel:	-		
Opmerkingen:	- Na de diverse testen en discussie M. Staring gekozen voor een algoritme met zowel een maskering in IR als GV. Op basis van second derivative zal na extra contrast enhancement en registration alsnog een transformix uitgevoerd moeten worden terwijl bij een maskering direct het resultaat wordt verkregen. Door ook gebruik te maken van een maskering in GV zal		

	mis-registration als gevolg van de horizontale “streep” in GV beeld worden voorkomen. Beide masks zijn op basis van bounding box van de gesegmenteerde persoon (d.m.v. threshold) en worden ge-dilated om er zeker van te zijn dat in beide gevallen de rand van de persoon meegenomen worden. De bounding box wordt verkregen door toepassing van labelling op de regions welke bij de segmentatie wordt verkregen. Door aanname van a priori knowledge waar de persoon zich bevindt kan op basis van een coördinaat op persoon de juiste regio geselecteerd worden. Voor dit onderzoek kan deze aanname gedaan worden, maar in een echte situatie moet rekening gehouden worden met misalignment van de laser. Een tweede aanname welke is gedaan is op basis van a priori knowledge van FOV en aantal pixels kan de daadwerkelijke grootte van het GV deel in het IR deel bepaald worden. Daarmee kan registratie beperkt worden tot translatie en rotatie.
Resultaten:	-

Week:	22	Datum:	05-06-2011 Thuis
Activiteiten:	- Final testing beeldregistratie: testen methode van algoritme 03-06-2011 - Uitwerken registratie algoritme in matlab		
Doel:	- T.b.v. oplossen registratie problemen en uitvoeren registratie		
Opmerkingen:	- Na afronden uitwerken algoritme, uitvoeren registratie alle scenario's		
Resultaten:	- Resultaat final test: OK, zowel rigid als affine transformation Links fixed mask, midden links moving mask, midden rechts resultaat, rechts resultaat op IR beeld. 		

Week:	23	Datum:	06-06-2011 @TNO
Activiteiten:	- Voortgangsbespreking met Piet - Bijwerken logboek a.d.h.v. voortgangsbespreking - Afronden en testen registratie algoritme		
Doel:	- T.b.v. registratie uitvoering		
Opmerkingen:	- Om registratie algoritme te kunnen runnen is er een coördinaat van het target in zowel IR als GV benodigd. Uitgangspunt onderzoek is dat target al is gevonden, dus dat deze coördinaten al bekend zijn. In reallife zal de applicatie dit dus moeten kunnen - Tevens is er een threshold input benodigd om een binary image te creëren met regions inclusief target. Threshold wordt handmatig bepaald in IR en GV images en zijn een waarde waarvan zeker is dat target als region wordt aangemerkt. - Size binary morph dilate operation op masks afhankelijk van de situatie		
Resultaten:	- Testen algoritme op scenario 1 en 2: algoritme werkt naar behoren - Thresholds zijn een nadeel, metname wanneer target nog beter verscholen is. Uitvoeren op andere scenarios moet dit uitwijzen		

Week:	23	Datum:	07-06-2011 Thuis
Activiteiten:	- Testen registratie algoritme andere scenarios - Werken aan presentatie student meeting		
Doel:	- T.b.v. registratie uitvoering en studentmeeting		
Opmerkingen:	-		
Resultaten:	- Testen algoritme op scenario 3: algoritme dient voor scenarios waarbij doel meer verscholen zit aangepast te worden. Aanpassingen: • histeq uitvoeren op IR, t.b.v. bounding box • morph. Oper. Uitvoeren op IR: imopen met square size 2 of meer (afhankelijk van situatie), daarnaast afhankelijk van de situatie vervolgens imdilate. Bij een van de scenarios ook imopen uitvoeren t.b.v. bepalen bounding box voor IR mask - Thresholds zijn een nadeel, metname wanneer target nog beter verscholen is. Uitvoeren op andere scenarios moet dit uitwijzen		

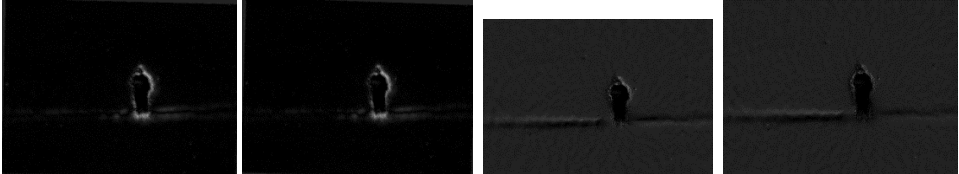
Week:	23	Datum:	08-06-2011 Thuis
Activiteiten:	- Adfronden presentatie student meeting		
Doel:	- T.b.v. studentmeeting en voortgangsmeting		
Opmerkingen:	-		
Resultaten:	- Presentatie status onderzoek		

Week:	23	Datum:	09-06-2011 @TU-Delft
Activiteiten:	- Presentatie student meeting - Voortgangsmeting		
Doel:	- T.b.v. studentmeeting en voortgang		
Opmerkingen:	Commentaar: - Groundtruth creëren voor beeldregistratie: bepalen wat nu daadwerkelijk de beste settings voor elastix zijn, daarna alles registreren - Goed verwoorden waarom zowel auto-recognition als door human. Tevens goed verwoorden waarom aannames zijn gedaan voor registratie (paragraaf weiden aan discussie) - Weight tussen IQM_{p1} en IQM_{p2} bepalen (p1 = patch 1 ofwel gehele image, p2 = patch 2 ofwel gebied rondom object) d.m.v. verificatie door human - Mogelijk toch uiteindelijk een video stream maken (toevoegen herkenning in een videostream i.p.v. stilstaande beelden) - Volgende keer: • Methodes in fusie toolboxes toelichten • IQM methode toelichten • Methode vaststellen IQM_{p1} x IQM_{p2} bepaald		
Resultaten:	-		

Week:	23	Datum:	12-06-2011 Thuis
Activiteiten:	- Final registration testing: • voor scenario 1 diverse registration results elastix bepaald voor verschillende parameter settings • 2 varianten groundtruth gecreëerd: exacte fit met deformatie target (is exacte fit, maar niet juiste situatie door gewijzigde aspect ratio op GV) en optimale fit zonder deformatie van target en met kleine scaling (theoretisch optimale fit) • Algoritme gemaakt: absdiff groundtruth met elastix resultaat met vervolgens mean hiervan bepaald en tevens correlation coëff groundtruth en elastix - Toevoegen tekst file output registratie algoritme t.b.v. fusie		
Doel:	- T.b.v. final registration		
Opmerkingen:	- Tabel gemaakt met varianten elastix settings met resultaten (appendix A)		
Resultaten:	- In eerste instantie lijkt: • Op basis van eigen waarneming absdiff: rigid transformatie met 3000 iterations (rigid variant 5) • Op basis correlatie: similarity met 3000 iterations (similarity variant 4)		

Week:	24	Datum:	13-06-2011 Thuis
Activiteiten:	- Final registration testing: • Extra variant groundtruth gecreëerd: betere optimale fit met meer scaling maar binnen edge met interpolatie artefacts IR upsampling (op het oog beste groundtruth) • Tekst output resultaten gemaakt • Histogram output absdiff gemaakt en std bijgevoegd - Tabel resultaten aangepast		
Doel:	-		
Opmerkingen:	- Tabel gemaakt met varianten elastix settings met resultaten (appendix A)		
Resultaten:	- Waarneming: • Op basis van eigen waarneming absdiff: similarity variant 3 • Op basis gemeente stats: similarity variant 4 of rigid variant 1		

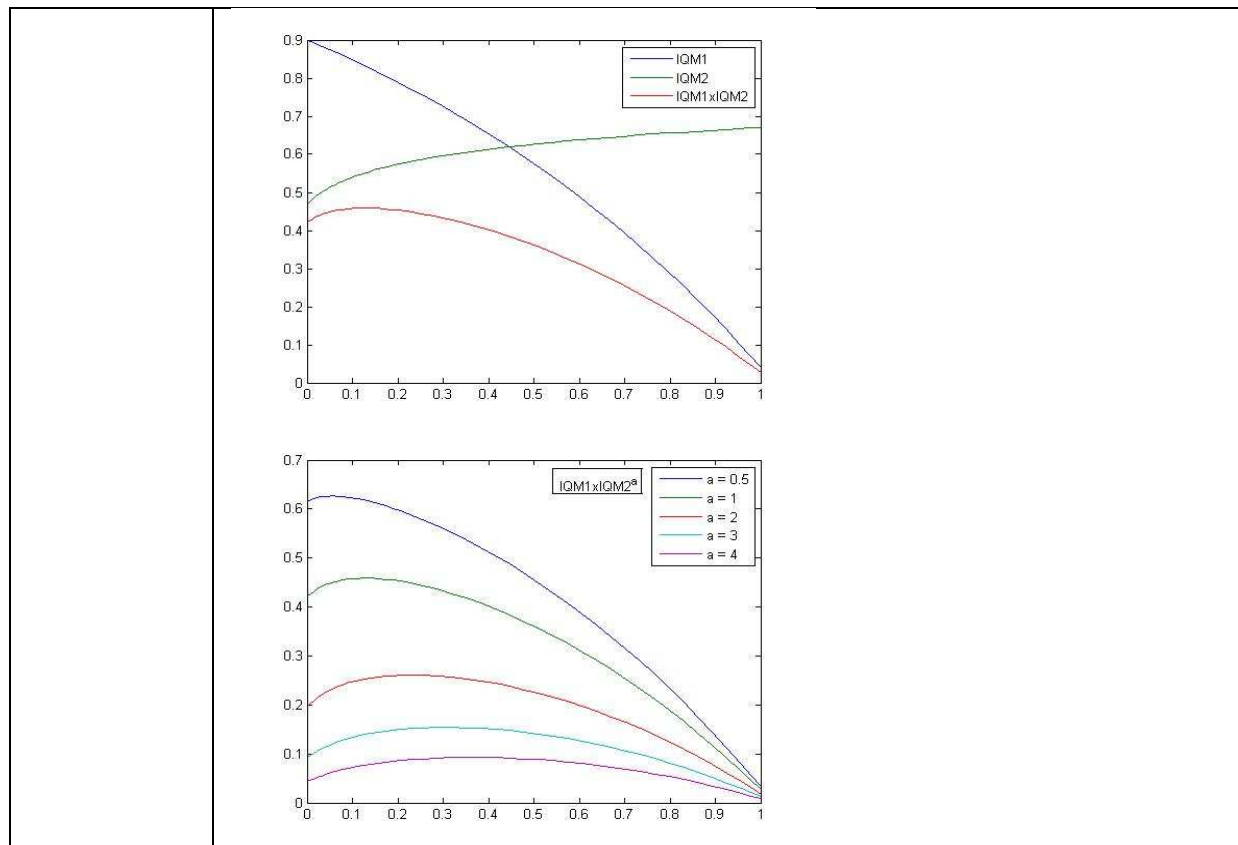
Week:	24	Datum:	14-06-2011 Thuis
Activiteiten:	- Final registration testing: • Nog een extra variant groundtruth gecreëerd: geen scaling (aangezien door meenemen FOV IR en GV in registratie algoritme een rigide transformatie aangenomen kan worden, zie week 19 15-05-2011) - Tabel resultaten aangepast - Scenario 1 t/m 3 doorlopen met registratie algoritme		
Doel:	- T.b.v. final registration		
Opmerkingen:	- Tabel zie onder (appendix A) - Aanname: opt2 is beste groundtruth (scaling maar binnen edge met interpolatie artefacts in upsampled IR image) ondanks meenemen FOV in		

	registratie algoritme (kleine hoeveelheid scaling door afrondingsfouten wel verwachtbaar). Waarom: mooiste waarneming met het menselijk oog
Resultaten:	- Enkele absdiff images met beste optimalfit als groundtruth links naar rechts de 4 beste resultaten: rigid 1, rigid 6, similarity 1 en similarity 4 (overige data in tabel appendix A). Noot: the mean value van similarity is aan de hoge kant, omdat bij elastix ruis weg is gefiltert (wel aanwezig in groundtruth). 

Week:	24	Datum:	19-06-2011 Thuis
Activiteiten:	- Scenario 4 t/m 7 doorlopen met registratie algoritme - Elastix voor alle 7 scenarios gerund op basis resultaat tabel		
Doel:	- T.b.v. final registration		
Opmerkingen:	- Scenarios zie onder appendix B (tevens alle matlab files: appendix C) - Controle nog uitvoeren!		
Resultaten:	- Zie hieronder (appendix B)		

Week:	25	Datum:	20-06-2011 @TNO
Activiteiten:	- Logboek bijwerken - IQM studie: keuze en opstellen optimalisatie algoritme met weight		
Doel:	- T.b.v. evaluatie		
Opmerkingen:	- Voorstel optimalisatie: $IQM_{p1} \times IQM_{p2}^{\alpha}$ en deze testen tegen andere mogelijkheden voor weight. Weight is n.a.v. de voortgansmeeting 9 juni (aanpassingen op fusie plan weken 15 t/m 18).		
Resultaten:	- Optimalisatie algoritme voorstel: $IQM_{p1} \times IQM_{p2}^{\alpha}$		

Week:	25	Datum:	21-06-2011 Thuis
Activiteiten:	- Controle registratie resultaten - Vervolg IQM optimalisatie algoritme (testen t.o.v. andere weight functies)		
Doel:	- T.b.v. final registration - T.b.v. evaluatie human recognition		
Opmerkingen:	- Registratie mogelijk 2 scenarios handmatige correctie: niet noodzakelijk, maar wel mooier uit esthetisch oogpunt		
Resultaten:	- Registratie gereed! - Optimalisatie algoritme $IQM_{p1} \times IQM_{p2}^{\alpha}$: • boven theoretisch voorbeeld IQM_{p1}, IQM_{p2} en $IQM_{p1} \times IQM_{p2}$; verticaal Image Quality en horizontaal de dimming ratio • onder $IQM_{p1} \times IQM_{p2}^{\alpha}$ voor hetzelfde voorbeeld voor verschillende waarden van α		



Week:	25	Datum:	26-06-2011 Thuis
Activiteiten:	- Vervolg IQM studie en opzet "human in the loop" (aanpassingen op fusie plan weken 15 t/m 18 n.a.v. voortgangsmeting 9 juni) - Keuze IQM		
Doel:	- T.b.v. evaluatie human recognition		
Opmerkingen:	- Human in the loop: • IQM eerst gebruiken op target region om te bepalen welke fusie algoritme de hoogste Quality heeft (eventueel bepalen voor variabele dimming ratio). • Daarna op de 'winnaar' de dimming ratio toepassen met optimalisatie algoritme $IQM_{p1} \times IQM_{p2}^a$, bepalen optimale α. • Human in the loop: bepalen α door inbreng evaluatie mensen. ▪ Groepjes van x personen in resultaat objecten laten classificeren: humanmade objects (vehicles {cars, trucks, tanks}, fences, buildings etc) of living creatures {human, animals}. Optimum $IQM_{p1} \times IQM_{p2}^a$ voor een x aantal scenarios per waarde voor α: een groep met oplopende α (starten met optimale background/scenery information) een groep met afnemende α (starten met optimale human information). Zo ontstaat een gebied tussen twee waarden voor α waarin alle objecten geclassificeerd kunnen worden en een gebied waarin de persoon geclassificeerd kan worden (per waarde voor α een classification rate voor		

	alles en voor alleen het target) ▪ Vervolgens ook vragen of iedereen de resultaten voor de verschillende α warden op volgorde beste tot slechtste plaatje te plaatsen. Hier een waarde tussen 0 en 1 aan koppelen (1 beste, 0 slechtste). Zo wordt een Quality rate bepaald. ▪ Classification rates vervolgens vermenigvuldigen met de Quality rate levert een waarde voor α welke optimaal is.
Resultaten:	- Keuze IQM: Piella metric, omdat deze zowel salient details alsmede edge informatie meeweegt en werkt met weight voor windows en per window tussen de twee te fuseren beelden; er wordt gebruik gemaakt van benadering HVS (Human Visual System)

Week:	26	Datum:	27-06-2011 @TNO
Activiteiten:	- Logboek bijwerken - Werkende krijgen van bestaande toolboxes (start met MATIFUS wavelet) - Average weight fusion toegepast (algoritme geschreven zonder PCA, diverse weights toegepast)		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	- Average weight fusion (onder links weight 0.5 op patch IR) 		

Week:	26	Datum:	03-07-2011 Thuis
Activiteiten:	- Werkende krijgen van bestaande toolboxes (start met MATIFUS wavelet)		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	- MATIFUS werkende gekregen		
Resultaten:	- Eerste fusie uitgevoerd: standaard wavelet (coif) en average fusie regel		

Week:	27	Datum:	10-07-2011 Thuis
Activiteiten:	- Werkende krijgen van bestaande toolboxes: • Lezen ICA toolbox paper • Lezen DT-CWT toolbox paper		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	-		

Week:	28	Datum:	11-07-2011 @TNO
Activiteiten:	- Bijwerken logboek - Uitbreiden average fusion met PCA regel		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	- PCA tb.v. fusie niet open source beschikbaar, dus zelf in matlab gemaakt		
Resultaten:	-		

Week:	28	Datum:	12-07-2011 @TNO
Activiteiten:	- Afronden PCA algoritme t.b.v. (average) weight fusion		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	-		

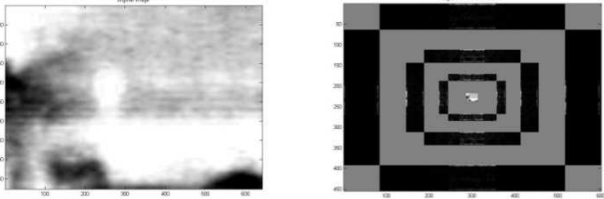
Week:	28	Datum:	17-07-2011 thuis
Activiteiten:	- Testen PCA regel t.b.v. (average) weight fusion		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	- Average weight fusion gereed		

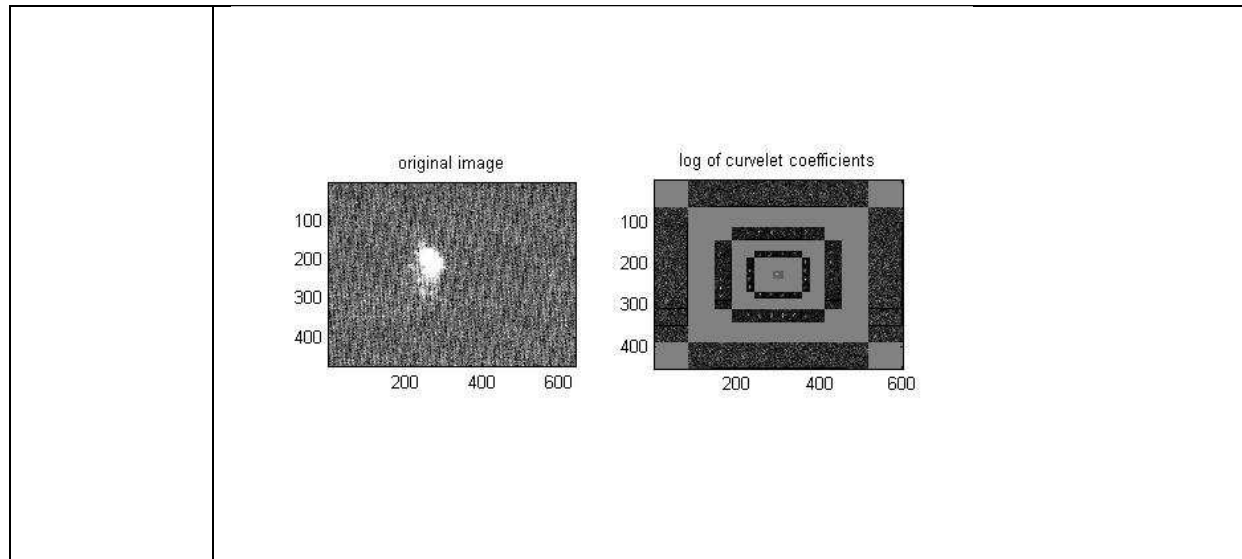
Week:	29	Datum:	18-07-2011 @TNO
Activiteiten:	- Bijwerken logboek - Uitwerken papers DT-CWT en inhoud toolbox		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	-		

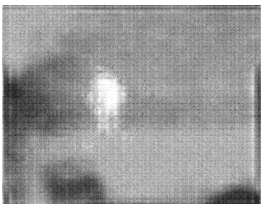
Week:	29	Datum:	19-07-2011 Thuis
Activiteiten:	- Werkende krijgen van bestaande toolboxes: • Werken aan basis fusie algoritme gebruikmakend van DT-CWT toolbox		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	-		

Week:	29	Datum:	24-07-2011 Thuis
Activiteiten:	- Werkende krijgen van bestaande toolboxes: • Werken aan basis fusie algoritme gebruikmakend van DT-CWT toolbox		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	- DT-CWT toolbox werkende gekregen: decompositie DT-CWT. Fusie algoritme op basis van deze toolbox nog mee bezig. - Resterende toolboxes: curvelet, contourlet, ICA		
Resultaten:			

Week:	30	Datum:	25-07-2011 @TNO
Activiteiten:	- Bijwerken logboek - Werkende krijgen van bestaande toolboxes: • DT-CWT algoritme gereed		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	- DT-CWT: bij average fusion voor alle coëfficiënten is het resultaat gelijk aan average fusion. Bij gebruik maxabs en maxavg fusion rules heeft noise in GV image negatieve invloed op resultaat (rare artefacts).		
Resultaten:	- Avera/average fused, average/maxabs fused en average/maxavg fused 		

Week:	30	Datum:	27-07-2011 Thuis
Activiteiten:	- Werkende krijgen van bestaande toolboxes: • Curvelab toolbox werkende gekregen		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	Voorbeeld decompositie curvelet: 		



Week:	31	Datum:	01-08-2011 @TNO
Activiteiten:	- Bijwerken logboek - Afronden curvelet fusie bouwsteen - Contourlet bouwsteen eveneens gemaakt		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	- bij average fusion voor alle coëfficiënten is het resultaat gelijk aan average fusion. Bij gebruik maxabs en maxavg fusion rules heeft noise in GV image negatieve invloed op resultaat (rare artefacts).		
Resultaten:	- Curvelet: Avera/average fused, average/maxabs fused en average/maxavg fused    - Contourlet: Avera/average fused, average/maxabs fused en average/maxavg fused   		

Week:	31	Datum:	02-08-2011 Thuis
Activiteiten:	- Werkende krijgen van bestaande toolboxes: • Werken aan basis fusie algoritme gebruikmakend van ICA		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	- ICA toolbox werkt alleen met MRI beelden (EEG, fMRI en sMRI). Nog niet werkende gekregen voor willekeurige extensies (tif, jpg, png etc) - Andere toolbox zoeken		
Resultaten:			

Week:	32	Datum:	08-08-2011 @TNO
Activiteiten:	- Bijwerken logboek - Zoeken nieuwe ICA toolbox - Programmeren Piella IQM		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	Wang & Bovik: $Q_0 = Q(x, y) = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\mu_x \mu_y}{\mu_x^2 + \mu_y^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}$ Piella: $Q(I_1, I_2, F) = \sum_{w \in W} c(w) (\lambda_w Q_0(I_1, F w) + (1 - \lambda_w) Q_0(I_2, F w))$ Piella edge: $Q = Q(I_1, I_2, F) \cdot Q(I_1', I_2', F')^\alpha$ Saliency: $\lambda = \sigma_{I_1}^2 / (\sigma_{I_1}^2 + \sigma_{I_2}^2)$		
Resultaten:			

Week:	32	Datum:	10-08-2011 Thuis
Activiteiten:	- Afronden programmeren en testen Piella IQM		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	Algoritme gereed, maar is langzaam. Nogmaals bekijken of deze efficiënter gemaakt kan worden		
Resultaten:			

Week:	33	Datum:	15-08-2011 @TNO
Activiteiten:	- Bijwerken logboek - testen en verbeteren Piella IQM - vaststellen hoe IQM te gebruiken		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	Nadelen Piella: - Bepaalt specifiek kwaliteit fusie adhv hoeveelheid "informatie" (salient details) van beide beelden in het fusiebeeld terecht zijn gekomen. Dus beide beelden worden in kwaliteit meegewogen terwijl in GV-IR geval er ook sprake is van gebieden met alleen IR of GV informatie. - Geen maat voor goede image. Qualitätsmaat op basis van information		

	preservation. In situatie IR-GV dient de kwaliteit ook een maat te zijn voor verbetering t.o.v. initiele situatie (alleen IR). Dus voegt GV wat toe. - Piella kijkt naar salient details en edges, maar niet naar background preservation. Oplossingen: - Alternatief ontwikkelen op basis van Piella waarbij f met IR wordt vergeleken - Groudtruth voor piella creëren: $Q(IR, GV, IR)$. Deze vergelijken met $Q(IR, GV, f)$.
Resultaten:	

Week:	33	Datum:	21-08-2011 Thuis
Activiteiten:	- Alternatief voor Piella IQM uitgewerkt		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	Alternatief: - General saliency op basis variance: $S = \frac{1}{N} \sum_{w=1}^N \frac{\sigma^2_{f w}}{\sigma^2_{f w} + \sigma^2_{IR w}}$ - Target extraction (edge recognition): $E = \frac{1}{N} \sum_{w=1}^N \frac{\sigma^2_{f' w}}{\sigma^2_{f' w} + \sigma^2_{IR' w}}$ en f' en IR' de gradient magnitude - Background preservation: $Q_{0(IR f)}$, hetgeen gedeelte van piella IQM is - Uiteindelijke metric: $Q = S \cdot E \cdot Q_{0(IR f)}$, mogelijk weging toevoegen - Indien S of E < 0.5: fusie resultaat is slechter dan initieel IR		
Resultaten:			

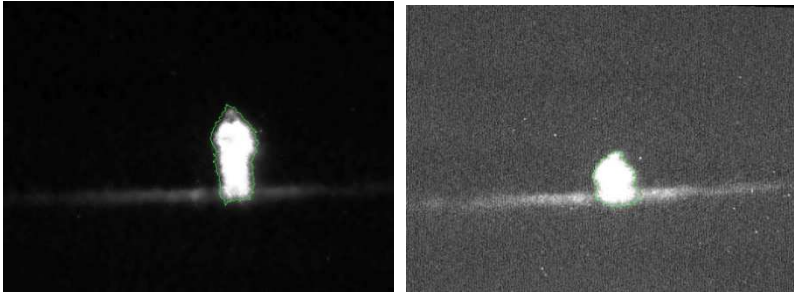
Week:	34	Datum:	22-08-2011 @TNO
Activiteiten:	- Bijwerken logboek - Schrijven fusie algoritme: basis fusie en segmentation		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:			

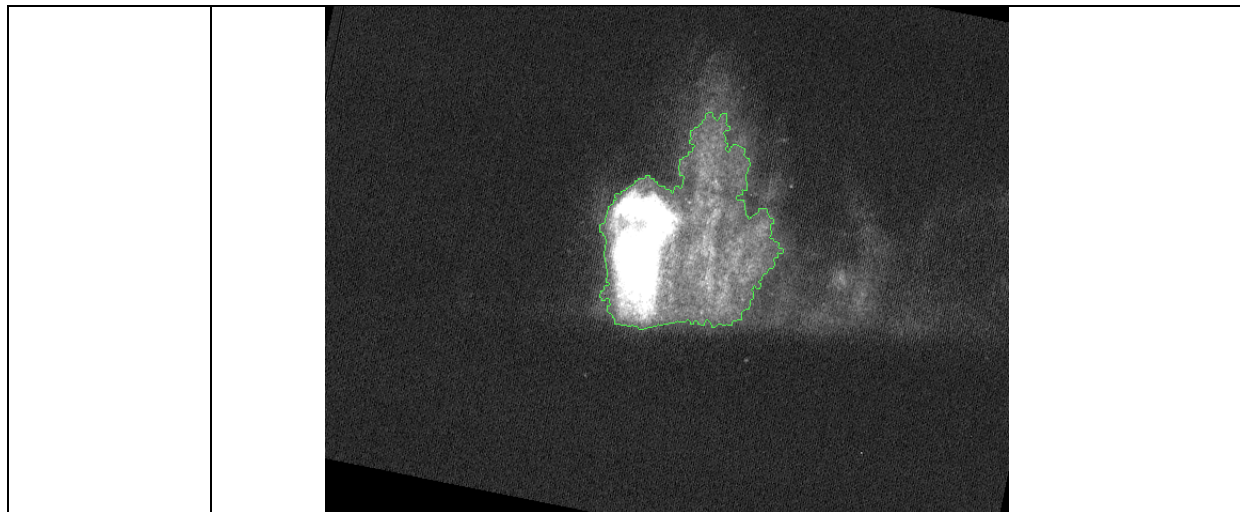
Week:	34	Datum:	23-08-2011 @Thuis
Activiteiten:	- Presentatie voortgangsmeting maken		
Doel:	- T.b.v. voortgangsmeting		
Opmerkingen:	-		
Resultaten:			

Week:	34	Datum:	26-08-2011 @TUD
Activiteiten:	- voortgangsmeting		
Doel:	- T.b.v. voortgangsmeting		
Opmerkingen:	- Volgende keer eerste resultaten - Volgende keer naast Emile en Piet eveneens Marco Loog uitnodigen		

	- M.b.t. Image Quality discussie met Ingrid Heynricks opstarten
Resultaten:	

Week:	34	Datum:	28-08-2011 @Thuis
Activiteiten:	- Segmentation algoritme maken t.b.v. region based fusion		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	- Segmentation mogelijkheden op basis morphological operations of gradient method gecombineerd met threshold en deze testen op alle scenarios. Voor alle scenarios een oplossing mogelijk behalve scenario 3. Wel zelf per scenario een keuze maken.		
Resultaten:			

Week:	35	Datum:	04-09-2011 @Thuis
Activiteiten:	- Segmentation algoritme maken t.b.v. region based fusion afmaken - Aanpassen IQM: gebieden met var = 0 kennen geen quality, daarom waarde 0 aanpassewn tot hele kleine waarde 0.0001. - IQM snellere loop gemaakt (op basis bestaande matlab algoritme blockprocessing was helaas niet mogelijk)		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	- Segmentation methods in een algoritme geplaatst met keuze parameters als input. Segmentation op beelden zichtbaar gemaakt (zie hieronder) - Waarde 0.0001 voor var in IQM bepaald d.m.v. testen met regel in algoritme if var == 0 then var = 0.xxxx: • 0.1: IQM = 0.4175 • 0.01: IQM = 0.4179 • 0.001: IQM = 0.4180 • 0.0001: IQM = 0.4180		
Resultaten:	- IQM algoritme gereed edoch i.v.m. vele loopjes duurt de berekening veel te lang. Dit later oplossen. - Segmentation gereede, behalve voor scenario 3. Hieronder 3 voorbeelden, scenario 1 op basis gradient (beste methode met weinig/geen ruis), scenario 2 op basis morphological operators en scenario 3 op basis morphological operators (beste resultaat) 		



Week:	38	Datum:	19-09-2011 @TNO
Activiteiten:	- Bijwerken logboek - Algoritme voor positionering fusiebeeld in initiele IR beeld gemaakt - Pixel based fusie algoritme afgerond (fusie totaalbeelden m.b.v. bouwstenen: zgn simpele bestaande fusie)		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	- Fusiealgoritme pixelbased fusion gereed, edoch verdient de fusieregel nog enige aandacht. Nu is het op basis van weighted average voor laagfrequente coëfficiënten en maxabs voor hoogfrequente coëfficiënten. Dit kan betekenen dat belangrijke features in laagfrequente coëfficiënten weg kunnen vallen en ruis in hoogfrequente coëfficiënten worden meegenomen. Een meer intelligente fusieregel op alle coëfficiënten zou dit mogelijk kunnen voorkomen.		
Resultaten:			

Week:	38	Datum:	25-09-2011 Thuis
Activiteiten:	- Vaststellen diverse mogelijke fusieregels - Start regionbased fusiealgoritme		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	- Priority fusion als extra methode toegevoegd: region 1 = GV en region 2 = IR op basis a priori kennis.		
Resultaten:			

Week:	39	Datum:	26-09-2011 @TNO
Activiteiten:	- Nieuwe werkplek - Discussie fusieregels - Discussie segmentation oplossing scenario 3		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:			
Week:	39	Datum:	30-09-2011 @TNO inhaaldag

Activiteiten:	- Nieuwe werkplek: internet en bureau - Fusieregels bepaald: local variance en local maxabs voor weigthed average en weighted average voor LF componenten, en PCA ook voor weighted average. - Sliding window algoritme t.b.v. fusieregels gemaakt
Doel:	- T.b.v. fusie algoritme
Opmerkingen:	- Keuze fusieregels gebaseerd op maximaliseren salient details en edges. Dit betekend dat noiseremoval noodzakelijk is. Volgende fusieregel mogelijkheden: • LF cf en weigthed average: ▪ Weigthed average met gewicht 0.5 (risico wegvallen details) ▪ Weigthed average met gewicht per pixel bepaald door local variance (risico noise invloed) ▪ Voor LF ook max local variance per pixel (risico noise invloed) ▪ Voor Weighted Average ook gewicht m.b.v. PCA (risico wegvallen van een van de beelden) • HF cf: ▪ Max abs value ▪ Max local abs value ▪ Eventueel max local variance (risico wegvallen details)
Resultaten:	

Week:	40	Datum:	03-10-2011 @TNO
Activiteiten:	- Weighted average met local variance en local maxabs fusion rule - PCA aangepast - Afspraak gemaakt TU-Delft met MMI t.b.v. IQM discussie		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:			

Week:	40	Datum:	05-10-2011 Thuis
Activiteiten:	- Start region based fusion algoritmes aanpassen		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:			

Week:	40	Datum:	09-10-2011 Thuis
Activiteiten:	- Region based fusion algoritmes aanpassen		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	- Weighted average en priority fusion gereed		

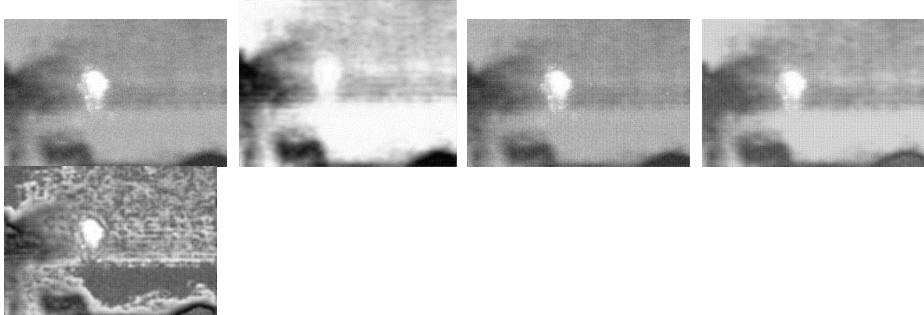
Week:	41	Datum:	10-10-2011 @TNO
Activiteiten:	- Region based fusion algoritmes aanpassen		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	- DT-CWT region based fusion gereed		

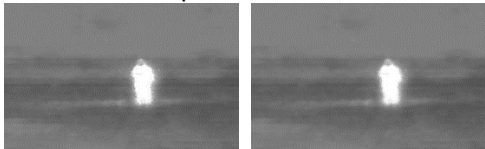
Week:	41	Datum:	11-10-2011 Thuis
Activiteiten:	- Region based fusion algoritmes aanpassen		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	- Curvelet region based fusion gereed		

Week:	41	Datum:	14-10-2011 @TU Delft
Activiteiten:	- Discussie IQM methoden		
Doel:	- T.b.v. evaluatie		
Opmerkingen:	- Nadeel Piella: information transfer based. We missen een methode welke human interpretatie meeneemt. Metname expert operator invloed is van belang. Hiervoor bestaat nog geen algoritme. Wel ontwikkeling bij TU-Delft op medisch gebied en eye tracker, maar nog geen bruikbare algoritme. Conclusie: toch kijken naar eerder stadium gebruik maken van human in the loop (true/false positives). Piella kan gebruikt worden met aangepaste saliency, maar testen met mensen om te kijken hoe goed Piella presteert. Bij dimmen van achtergrond werkt Piella niet voor achtergrond, omdat dit non-structural changes zijn (informatie blijft gelijk, alleen contrast neemt af). Dus hier human input belangrijk. Wat nu doen: • Piella gebruiken met aangepaste saliency maat: niet variance, maar gebaseerd op taak operator (zoeken geschikte methode) ▪ Kijk naar belangrijke herkenningspunten en baseer saliency hierop • Human in the loop (experts) gebruiken voor testen hoe Piella presteert • Bij achtergronddimmen niet alleen Piella voor achtergrondmaat. Zoek een methode die contrast meet.		
Resultaten:	-		

Week:	41	Datum:	16-10-2011 Thuis
Activiteiten:	- Region based fusion algoritmes aanpassen		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	Nog doen: - Ruisreductie - Artefacts bij fusieregels bekijken - Afronden gecombineerde matlab code - Saliency maat zoeken - Prepareren beelden t.b.v. runnen algoritme		
Resultaten:	- Contourlet region based fusion gereed		

Week:	42	Datum:	17-10-2011 @TNO
Activiteiten:	- Region based fusion algoritmes aanpassen - Voorbereiden beelden - Runnen tests algoritmes		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	- Algoritme functioneert. Nu resultaten analyseren om te kijken of algoritme ook goed werkt		
Resultaten:	- Gecombineerde Fusie matlab code - Algoritme functioneert		

Week:	42	Datum:	19-10-2011 Thuis
Activiteiten:	- Region based fusion testen/runnen		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	N.a.v. resultaten op scenario 7 - Ruisreductie nog doorvoeren vanwege negatieve invloed op variance methoden - Kijken naar fusie regels. Eerste selectie op basis visuele observatie, niet op basis metric: welke werken niet goed, welke wel goed, welke aanpassen of weglaten, of nieuwe toevoegen		
Resultaten:	- Eerste resultaten scenario 1, hieronder de slechte resultaten weergegeven: links naar rechts 0.5 weighted average, contourlet maxabs, curvelet maxabs, DT-CWT maxabs, en local variance based (hier met DT-CWT, maar anderen niet beter) 		

Week:	43	Datum:	24-10-2011 @TNO
Activiteiten:	- Logboek bijwerken - Correctives doorvoeren op basis van resultaten: artefacts weggewerkt (juiste locatie absolute values in algoritme) - Ruisreductie: wiener filter op weighted average en threshold in wavelet		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	- Tests op scenario 1: verbeterde DT-CWT maxabs en contourlet maxabs 		

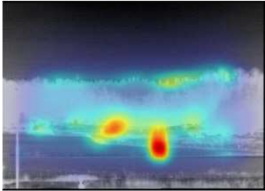
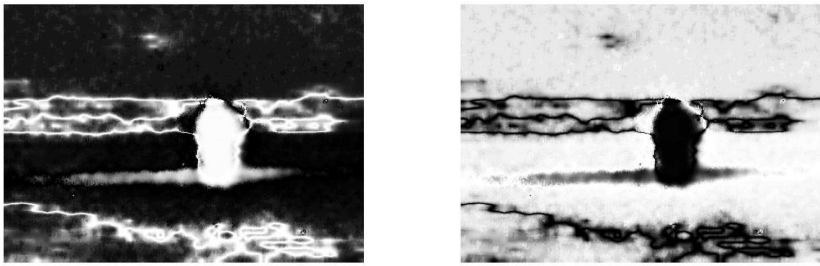
Week:	43	Datum:	25-10-2011 thuis
Activiteiten:	- Ruisreductie: threshold ook in curvelet en contourlet - Voorbereiden alle scenarios t.b.v. runnen algoritme		
Doel:	- T.b.v. fusie algoritme		
Opmerkingen:	-		
Resultaten:	-		

Week:	43	Datum:	27-10-2011 @TU-Delft
Activiteiten:	- studentmeeting		
Doel:	-		
Opmerkingen:	-		
Resultaten:	-		

Week:	44	Datum:	31-10-2011 @TNO
Activiteiten:	- Runnen algoritme op scenario 1 - Visuele evaluatie en deselectie fusieregels welke slechte resultaten geven		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	Volgende fusieregels afgekeurd: - Local variance op LF (alleen pixelbased) - max loc variance op LF (alleen pixelbased) - max abs op LF - Max average op LF - Weighted average met weight op basis local variance Goed resultaat: - Pixelbased met local max op LF: buiten doel metname IR - Region based Pixelbased over het algemeen donkerder gebied op locatie GV. Possitief voor focussen op doel, maar negatief voor afleiding van omgeving (slecht voor behoud achtergrond)		
Resultaten:	-		

Week:	44	Datum:	06-11-2011 thuis
Activiteiten:	- Studie saliency map		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	Paper: - Toet, Computational versus psychophysical bottom-up image saliency: a comparative evaluation study - Itti et al, a model of saliency-based visual attention for rapid scene analysis		
Resultaten:	-		

Week:	45	Datum:	07-11-2011 @TNO
Activiteiten:	- Studie saliency map		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	- Op basis literatuur studie en beschikbare modellen keuze saliencymap - Geschikt (edoch metname voor kleurenbeelden): Graph-based visual saliency gebaseerd (GBVS) op Itti et al (meer op edge beelden) en Frequency-tuned saliency (FTS)		

	- Getest in piella: vervangen variance door saliency map
Resultaten:	Saliency maps toepasbaar in IQM (Piella) Saliencymap op IR cf Itti-Koch:  Weight voor saliency in piella op basis FTS $S(x, y) = \ I_\mu - I_G\$, links voor GV en rechts het corresponderende weight voor IR 

Week:	45	Datum:	08-11-2011 thuis
Activiteiten:	- Aanpassingen IQM, op basis sliding window i.p.v. loopjes (loopjes nemen veel rekentijd in beslag)		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	- Saliencymap toegevoegd		
Resultaten:	-		

Week:	45	Datum:	13-11-2011 thuis
Activiteiten:	- Verder met aanpassingen IQM, op basis sliding window		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	- Nog op te lossen probleem: covariance tussen twee windows		
Resultaten:	-		
Week:	47	Datum:	21-11-2011 TNO
Activiteiten:	- Oplossen probleem IQM		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	- Nog niet gelukt. Probleem sliding window verholpen, maar resultaat is niet correct (IQM zeer klein, matcht niet met IQM middels loopfuncties).		
Resultaten:	-		

Week:	47	Datum:	22-11-2011 thuis
Activiteiten:	- Voorbereiden voortgangsmeting		
Doel:	- Voortgangsbewaking		
Opmerkingen:	-		
Resultaten:	-		

Week:	47	Datum:	24-11-2011 @TU-Delft
Activiteiten:	- Voortgangsmeting		
Doel:	- Voortgangbewaking		
Opmerkingen:	- Eerst draft paper herkenning door de mens aanleveren: januari - Kijken naar smoothing i.p.v. FTS, tenzij FTS werkt - Ook Harris corners toepassen als saliency maat - Kijken of task dependency ook implementeerbaar is voor saliency - Bij dimming: kan ook niet-lineair, bijvoorbeeld local		
Resultaten:	-		

Week:	47	Datum:	27-11-2011 thuis
Activiteiten:	- Voorbereiden uitzetten human in the loop - Verbeteren IQM		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	- Verbeteren IQM: op internet zoeken naar oplossing probleem loopjes		
Resultaten:	- Eigenlijke implementatie Wang en Bovic gevonden		

Week:	48	Datum:	28-11-2011 @TNO
Activiteiten:	- Voorbereiden uitzetten human in the loop - Verbeteren IQM: hergebruiken Wang en Bovic en aanpassen naar Piella - Testen IQM		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	-		
Resultaten:	-		

Week:	48	Datum:	04-12-2011 Thuis
Activiteiten:	- Voorbereiden uitzetten human in the loop - Implementeren Saliency in IQM		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	-		
Resultaten:	-		

Week:	48	Datum:	05-12-2011 @TNO
Activiteiten:	- Uitzetten human in the loop: bij TNO Soesterberg - Download paper torralba en Navalpakkam/Itti - Uitwerken algoritme paper Marco Loog: harris corners t.b.v saliency		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	Human in the loop t.b.v. valideren IQM - TNO Soesterberg runt een menselijk experiment: experiment met proefpersonen waarbij beelden geordend worden met een rangorde. Het resultaat is dan een gemiddelde rangorde per beeldmodaliteit: januari 2012 - ik ga met de uitkomst van het experiment de Image Quality Algoritme evalueren. Dat kan met een correlatie tussen IQM waardes en de gegeven rangorde: januari 2012 - TNO Soesterberg gaat vervolgens de betrouwbaarheid van de resultaten van het experiment verifiëren d.m.v. meten		

	waarnemersconsistentie/reproduceerbaarheid: februari 2012
Resultaten:	-

Week:	49	Datum:	12-12-2011 @TNO
Activiteiten:	- Implementeren Harriscorner saliency - Implementeren Itti-Koch saliency - Runnen IQM op beelden welke uitgezet zijn voor human in the loop validatie		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	Harris algoritme: $S(x, y) = H(I) = \det T(I) - \kappa \cdot \text{trace}^2 T(I)$		
Resultaten:	- Zie bijlage C - De piella met toevoeging gradient beelden lijkt onzinnige resultaten te geven. Combinatie saliencymap en gradient lijkt ook dubbelop.		

Week:	1	Datum:	02-01-2012 Thuis
Activiteiten:	- Literatuurstudie task dependent saliency		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	- Paper Navalpakkam en Itti		
Resultaten:	-		

Week:	1	Datum:	05-01-2012 Thuis
Activiteiten:	- Studentmeeting bijwonen		
Doel:	-		
Opmerkingen:	-		
Resultaten:	-		

Week:	1	Datum:	08-01-2012 Thuis
Activiteiten:	- Literatuurstudie task dependent saliency		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	- Paper Torralba		
Resultaten:	-		

Week:	2	Datum:	09-01-2012 @TNO
Activiteiten:	- Bijwerken logbook - Paper torralba en Navalpakkam/Itti bestuderen en implementatie zoeken - Werken aan paper		
Doel:	- Evaluatie fusiealgoritme: task dependency		
Opmerkingen:	- Momenteel geen bestaande implementatie te verkrijgen. Dus voorlopig alleen bottom-up saliency gebruiken		
Resultaten:	-		

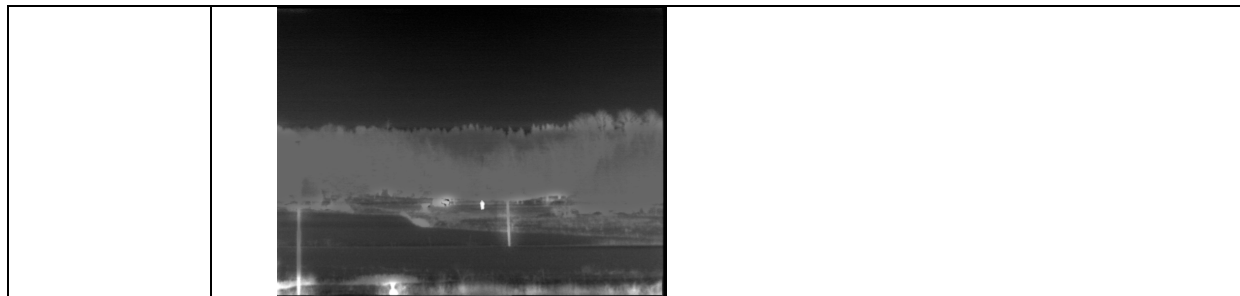
Week:	2	Datum:	15-01-2012 Thuis
Activiteiten:	- Werken aan paper		
Doel:	- paper		
Opmerkingen:	-		
Resultaten:	-		

Week:	3	Datum:	16-01-2012 @TNO
Activiteiten:	- Dimming algoritme - Start uitwerken algoritme voor video - Algoritme harris corner implementeren/verbeteren		
Doel:	- Achtergrond dimming - Fusie evaluatie		
Opmerkingen:	- 3 soorten dimming: global (gehele beeld behalve persoon), local (alleen rand rondom persoon) en contextual (lineair dimmen met saliency map als maat voor dimming) Global en local: $F_d = d \cdot kF + (1-k)F$ Contextual: $F_d = (d + (1-d)S) \cdot kF + (1-k)F$ met saliency map.		
Resultaten:	- Initiele dimming algoritme gereed maar nog niet getest		

Week:	3	Datum:	17-01-2012 Thuis
Activiteiten:	- Testen dimming algoritme en verbeteren		
Doel:	- Achtergrond dimming		
Opmerkingen:	- Rand effect: door dimming wordt de rand van de persoon zeer "onnatuurlijk" scherp. Smoothing toepassen (Gaussian). Nog niet geïmplementeerd.		
Resultaten:	-		

Week:	3	Datum:	22-01-2012 Thuis
Activiteiten:	- Werken aan paper - Verbeteren dimming (randeffect)		
Doel:	- Paper - dimmen		
Opmerkingen:	-		
Resultaten:	-		

Week:	4	Datum:	23-01-2012 @TNO
Activiteiten:	- Dimming afronden (randeffect, contextual op basis harris) - Dimming video maken		
Doel:	- Dimming		
Opmerkingen:	-		
Resultaten:	Voorbeeld contextual dimming (op basis harris):		



Week:	4	Datum:	29-01-2012 Thuis
Activiteiten:	- Berekenen IQM - Verbeteringen IQM (Saliency voor edge representatie local variance) - Opnieuw berekenen IQM - Transformix op alle frames voor video gedaan (registratie alle GV frames)		
Doel:	- Evaluatie fusiealgoritme		
Opmerkingen:	-		
Resultaten:	- Registratie video gereed - IQM gereed en waarden bepaald voor beelden welke in human experiment gedaan worden		

Week:	5	Datum:	30-01-2012 @TNO
Activiteiten:	- Bekijken resultaten human experiment en vergelijken IQM - Overleg Piet: hoe nu verder - Correlatie uitrekenen tussen IQM en human experiment		
Doel:	- Evaluatie fusiealgoritme - dimming		
Opmerkingen:	- Resultaten IQM in vergelijking met humanexperiment vertonen verschillen - Human experiment vindt pixelbase beste, IQM region based. Dit is logisch want region based geeft wel optimale informatieoverdracht en pixelbased geeft goede herkenning door mens. Afzonderlijk region based en pixel based bekeken zijn er wel overeenkomsten tussen IQM en human resultaten. Verder is quality van IR soms aardig voor IQM. Ook logisch aangezien IR grootste deel resultaat bevat. Derhalve beter om IR IQM waarde te negeren bij correlatie - Bespreking Piet: op basis bovenstaande IQM afzonderlijk correleren met human experiment: deel region based en deel pixelbased. Dan voor beide de beste fusie algoritme kiezen en alle scenarios fusie uitvoeren. Dan dimming uitvoeren, global, local en contextual voor beide in scenario met goed versholten persoon. Contextual dimming met saliency metric welke ook beste IQM geven. Dan wederom human experiment: beste dimming ratio per dimming algoritme en voor zowel pixel als region based. Daarna beste overall. Dan wederom vergelijken met IQM en vervolgens toepassen bewegende beeld: komt hier dezelfde mate van dimmen uit. Verder nog tot de conclusie gekomen dat dimming region based eigenlijk alleen global en contextual moet zijn en dimming zodra achtergrond ongeveer evenveel als blok rondom persoon gedimd is met gelijke dimming verder dimmen. Deze dus gaan implementeren.		

Resultaten:	- Uiteindelijke correlatie resultaat: - Over alles (correlatie IQM-Human alle beelden) beste resultaat: 0.8437, Wang-Bovik met Harris2 (harris2 = implementatie marco) - Over alles zonder IR: 0.7746, Wang-Bovik met harris2 - Region based zonder IR: 0.9509, Wang-Bovik met Harris (eigen implementatie, implementatie Marco follow-up met 0.9480) - Pixel based zonder IR: 0.9012, Piella met Harris2 - Overall best: wang-bovik met harris2 (altijd >0.7) - Pixel based: over het algemeen Piella beter - Region based: altijd wang-bovik beter
-------------	---

Week:	5	Datum:	31-01-2012 Thuis
Activiteiten:	- Werken aan dimming voor pixelbased		
Doel:	- dimmen		
Opmerkingen:	- alleen global dimming gereed		
Resultaten:	- global dimming voor region based		

Week:	5	Datum:	05-02-2012 Thuis
Activiteiten:	- Afronden dimming pixelbased		
Doel:	- Dimmen en evaluatie		
Opmerkingen:	-		
Resultaten:	- Dimming algoritme gereed		

Week:	6	Datum:	06-02-2012 @TNO en thuis
Activiteiten:	- Logboek bijwerken - Uitvoeren IQM op scenario 1 - Selectie beste fusie methoden		
Doel:	- Evaluatie fusie methoden		
Opmerkingen:	-		
Resultaten:	- Dimming algoritme gereed		

Week:	6	Datum:	12-02-2012 Thuis
Activiteiten:	- Runnen fusie scenario 2 t/m 7 - Runnen IQM scenario 2 t/m 7		
Doel:	- Evaluatie fusie methoden		
Opmerkingen:	-		
Resultaten:	-		

Week:	7	Datum:	13-02-2012 @TNO
Activiteiten:	- Ranking IQM scenario 1 t/m 7 (zie appendix D)		
Doel:	- Evaluatie fusie methoden		
Opmerkingen:	- Na uitvoeren fusie alle scenarios en ranking van de IQM waarden valt op dat bij regionbased fusion veel resultaten van de multi-resolution fusie methoden gelijke IQM waarden hebben.		
Resultaten:	- Winnaar pixelbased fusion: weighted average fusion met PCA. De overige methoden liggen vrij dicht bij elkaar qua ranking. Wel zijn voor zowel DT-		

	CWT en curvelet de beste resultatebn met HF-fusie regel maxlocabs. - Winnaar regionbased fusion: gedeeld winnaars 3x curvelet met LF fusieregel maxlocvar en priority fusion
--	---

Week:	7	Datum:	14-02-2012 Thuis
Activiteiten:	- Presentatie studentmeeting voorbereiden		
Doel:	- Studentmeeting		
Opmerkingen:	-		
Resultaten:	-		

Week:	7	Datum:	15-02-2012 Thuis
Activiteiten:	- Presentatie studentmeeting voorbereiden		
Doel:	- Studentmeeting		
Opmerkingen:	-		
Resultaten:	-		

Week:	7	Datum:	16-02-2012 Thuis
Activiteiten:	- Presentatie studentmeeting		
Doel:	- Studentmeeting		
Opmerkingen:	-		
Resultaten:	-		

Week:	7	Datum:	19-02-2012 Thuis
Activiteiten:	- Werken aan paper		
Doel:	- paper		
Opmerkingen:	-		
Resultaten:	-		

Week:	8	Datum:	20-02-2012 @TNO
Activiteiten:	- Werken aan paper - Discussie auto-recognition		
Doel:	- Paper - Voortgang en planning (wetenschappelijke relevantie auto-recognition)		
Opmerkingen:	- Vastgesteld dat auto-recognition in deze studie niet relevant is. Het is duidelijk dat auto herkenning in GV eenvoudig is en altijd verbetering oplevert. Fusie daarentegen op feature niveau zal geen verbetering opleveren t.o.v. classificatie in de beelden los van elkaar, vooral in GV. Daarbij geldt dat er alleen GV beelden van de persoon zijn. Dus herkenning van de mens zal alleen getrained kunnen worden op de mens zelf met 100% performance als resultaat. Ook valt het IR beeld zo te tweaken dat herkenning mogelijk zal zijn, maar weegt niet op tegen de mogelijkheden in GV. Kortom: niet relevant.		
Resultaten:	-		

Week:	8	Datum:	21-02-2012 Thuis
Activiteiten:	- Werken aan paper		
Doel:	- paper		
Opmerkingen:	-		
Resultaten:	-		

Week:	8	Datum:	23-02-2012 Thuis
Activiteiten:	- Voorbereiden voortgangsmeting		
Doel:	- voortgang		
Opmerkingen:	-		
Resultaten:	-		

Week:	8	Datum:	24-02-2012 @TU-Delft
Activiteiten:	- Voortgangsmeting		
Doel:	- voortgang		
Opmerkingen:	- Vastgesteld dat inderdaad autorecognition niet relevant en dus niet nodig is - Afmaken dimming en videostream met menselijke experimenten: dan is het waarschijnlijk wetenschappelijk voldoende - Paper belangrijkste, rapport is extra - Kijk nog even naar ranking correlatie bij bepalen optimale fusie methode		
Resultaten:	-		

Week:	8	Datum:	26-02-2012 Thuis
Activiteiten:	- Werken aan paper		
Doel:	- paper		
Opmerkingen:	-		
Resultaten:	-		

Week:	9	Datum:	27-02-2012 @TNO en thuis
Activiteiten:	- Werken aan paper		
Doel:	- paper		
Opmerkingen:	-		
Resultaten:	-		

Week:	9	Datum:	04-03-2012 Thuis
Activiteiten:	- Werken aan paper		
Doel:	- paper		
Opmerkingen:	-		
Resultaten:	-		

Week:	10	Datum:	11-03-2012 Thuis
Activiteiten:	- Werken aan paper - Werken aan IQM dimfit algoritme		
Doel:	- paper		
Opmerkingen:	-		
Resultaten:	-		

Week:	11	Datum:	12-03-2012 @TNO
Activiteiten:	- Uitvoeren Spearman rank correlation op fusie resultaten - Bekijken resultaten 2^e experiment		
Doel:	- Controle correlatie fusieresultaten en IQM fit op human experiment scenario 1 - Dimming experiment		
Opmerkingen:	-		
Resultaten:	- Ranking corelatie zie bijlage C: ranking correlatie geeft interessante inzichten, maar wijzigt resultaten niet. Ranking correlatie geeft gelijke correlatiewaarde voor meerdere saliency metrics. Voor pixelbased hoger, maar regionbased lager dan over alles. De gekozen methode op basis van eigenlijke correlatie zit tussen de hoogste ranking correlatie waardes, dus geen wijziging van IQM noodzakelijk. Verder geldt dat voor regionbased de ranking correlatie lager is, maar zowel bij regionbased afzonderlijk als voor alle resultaten samen geeft harris de beste correlatie. Voorkeur is toch voor eigenlijke correlatie omdat het een gemiddelde ranking is en niet iedereen in het experiment dezelfde ranking geeft. Het gemiddelde geeft dus niet de preciese ranking weer. Verder geeft de correlatie uiteindelijk de doorslag welke saliency metric echt de beste is. - 2^e experiment zie bijlage E: De wijze van opzet van het 2^e (dimming) experiment geeft goed inzicht op voorkeuren voor hoeveelheid dimming. Daarnaast biedt het een goede manier om op de overall best (volgens 6 personen) een IQM fit te doen, omdat voor deze dimming methode in de eerste ronde alle 16 personen een optimale hoeveelheid dimming hebben aangegeven. De distributies over contextual en global dimming zijn dusdanig mooi (met een duidelijk zichtbare piek) dat een keuze van optimale dimming mogelijk is en dat IQM fit mogelijk zou moeten zijn. Voor local dimming is de distributie te verspreid en is geen optimale dimming aan te wijzen en dus ook geen IQM fit mogelijk. Voor deze laatste zou dus een grotere groep voor een experiment genomen moeten worden om te testen of dan wel een piek in de distributie zichtbaar wordt. Verder is mooi zichtbaar dat de piek naar links verschuift als je van global naar contextual dimming gaat: dus bij contextual dimming kun je de achtergrond iets meer dimmen.		

Week:	11	Datum:	11-03-2012 Thuis
Activiteiten:	- IQM dimfit algoritme verbeteren en teopassen		
Doel:	- IQM fit voor dimming experiment		
Opmerkingen:	- Dimfit algoritme aangepast: W&B voor IR met fusiebeeld werkt (meer dimming levert lagere kwaliteit op), maar IQM op targetpatch werkt niet		

	(meer dimmen levert ook slechtere kwaliteit op). Dus dimmen levert slechtere kwaliteit. Dit is logisch want de IQM is ook gebruikt voor bepalen optimale fusie, dus iedere wijziging levert slechter resultaat op. Dus andere oplossing voor IQM op target benodigd. Oplossing: target background contrast (Michelson contrast of visibility): $I_{\max} - I_{\min} / I_{\max} + I_{\min}$. Voor $I_{\max}$ is de mean van het tgt genomen en voor $I_{\min}$ de mean van de achtergrond. Dit levert dus een hogere waarde voor meer dimming. Dim IQM is nu: $Q_d = Q(IR, F_d) \cdot \left(\frac{\overline{F_{d_tgt}} - \overline{F_{d_bg}}}{\overline{F_{d_tgt}} + \overline{F_{d_bg}}} \right)^a$
Resultaten:	- Zie bijlage E

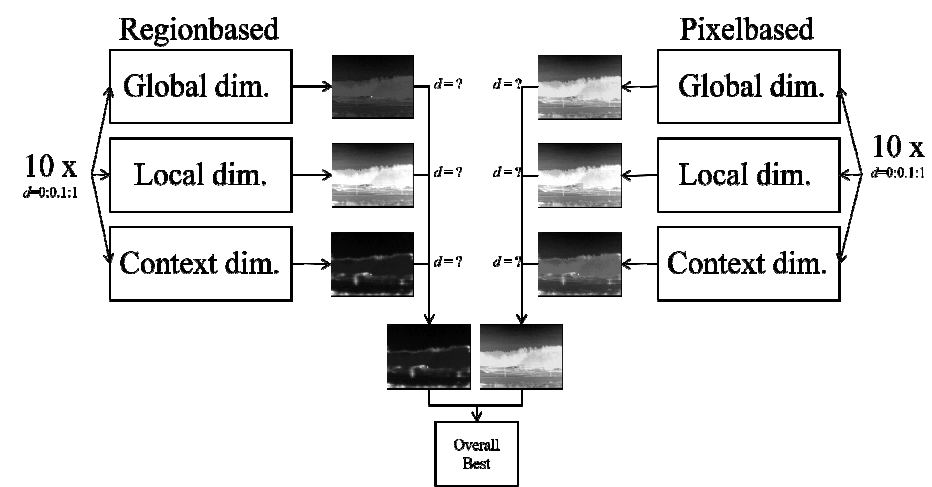
Week:	11	Datum:	15-03-2012 Thuis
Activiteiten:	- Student meeting		
Doel:	- Student meeting		
Opmerkingen:	-		
Resultaten:	-		

Week:	11	Datum:	18-03-2012 Thuis				
Activiteiten:	- Uitvoeren IQM dimfit						
Doel:	-						
Opmerkingen:	- Zie plaatjes bijlage E						
Resultaten:	Steering parameter	0.4	0.5	0.6	0.7	0.8	
	Pix_context	V	V	B	B	V	
	Pix_global	V	B	V	X	X	
	Reg_context	V	V	B	V	V	
	Reg_global	B	V	X	X	X	
	B = best fit, V = maximum bij juiste hoeveelheid dimming, X = geen fit Dus voor steering parameter waarde van 0.5 geeft de IQMfit voor alle dimmogelijkheden een bruikbare IQM (juiste optimale hoeveelheid dimming). Echter de beste fit (dus behalve maximum ook naastgelegen waardes correct) is steeds verschillend. Bij opsplitsing in context en global is er een bestfit parameter waarde van 0.6 te vinden voor context dimming en lijkt 0.4 of 0.5 de meest optimale voor global dimming. Opsplitsing in pixel en regionbased is minder succesvol.						

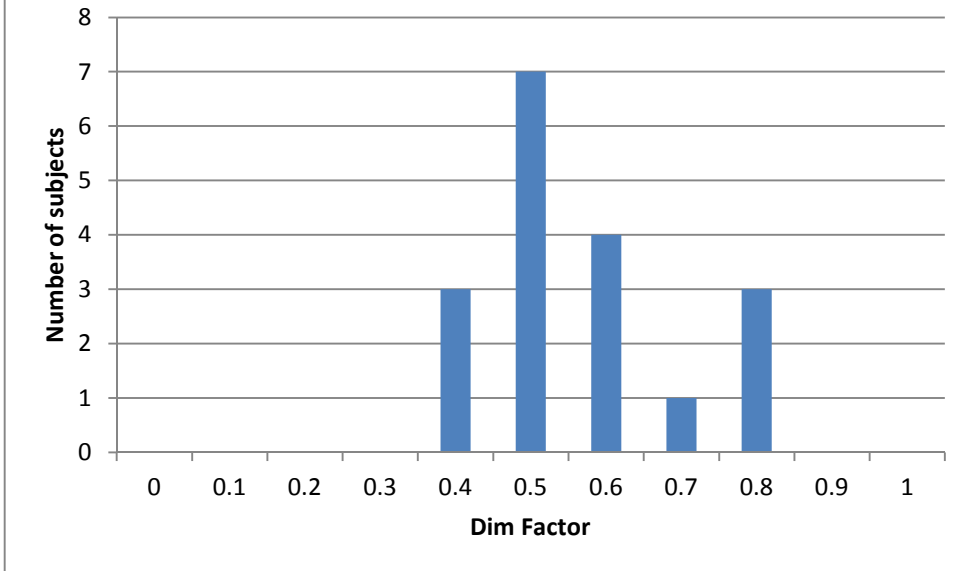
Week:	12	Datum:	19-03-2012 @TNO
Activiteiten:	- Bijwerken logboek - Controle resultaten 2^e experiment en IQM-fit - Maken videostreams 3^e experiment		
Doel:	- Afronden onderzoek fase		
Opmerkingen:	- nog aanpassingen benodigd in algoritmieken: • harrisalgoritme: aanpassing contrast verbetering (invloed complexer beeld)		

	- dimming: aanpassing voor dimfactor patch rondom persoon (meenemen harris context i.v.m. karakter scenario) - PCA geeft per video frame andere resultaten (invloed ruis). Daardoor flikkerend beeld in video. Daarom alleen bij 1^e frame PCA uitrekenen en deze voor alle frames gebruiken. Voor daadwerkelijk gebruik dient dus bij aanpassingen camera settings en veranderende lichtomstandigheden de PCA opnieuw bepaald te worden. Multi-resolutie methoden kennen dit probleem niet.
Resultaten:	- algoritme maken video stream gereed en werkt correct

Week:	12	Datum:	21-03-2012 Thuis
Activiteiten:	- Maken videostreams 3 ^e experiment		
Doel:	- Afronden onderzoek fase		
Opmerkingen:	-		
Resultaten:	- Alle aanpassingen doorgevoerd, videostreams gereed (11 stuks met dimfactor [0:0.1:1])		

Week:	13	Datum:	26-03-2012 @TNO
Activiteiten:	- Bijwerken logboek - uitzetten 3^e experiment - werken aan paper		
Doel:	- Afronden onderzoek fase		
Opmerkingen:	3^e experiment drie fases: - selectie dim factor voor alle dim methodes voor zowel pixel als region based fusion - selectie beste dim methode en dim factor per fusie methode - selectie overall best di methode en dim factor Zo kan inzicht verkregen worden over alle voorkeuren van de testpersonen. 		
Resultaten:	-		

Week:	14	Datum:	02-04-2012 @TNO
Activiteiten:	- Werken aan paper		
Doel:	- paper		
Opmerkingen:	-		
Resultaten:	-		

Week:	14	Datum:	09-04-2012 @TNO												
Activiteiten:	- Verwerken resultaten van 3 ^e experiment														
Doel:	- Afronden onderzoek														
Opmerkingen:	- Meer dimming dan statische frames - Maar ook grotere spreiding														
Resultaten:	 <table><caption>Histogram Data</caption><thead><tr><th>Dim Factor</th><th>Number of subjects</th></tr></thead><tbody><tr><td>0.4</td><td>3</td></tr><tr><td>0.5</td><td>7</td></tr><tr><td>0.6</td><td>4</td></tr><tr><td>0.7</td><td>1</td></tr><tr><td>0.8</td><td>3</td></tr></tbody></table>			Dim Factor	Number of subjects	0.4	3	0.5	7	0.6	4	0.7	1	0.8	3
Dim Factor	Number of subjects														
0.4	3														
0.5	7														
0.6	4														
0.7	1														
0.8	3														

Appendix A: Tabel registration results

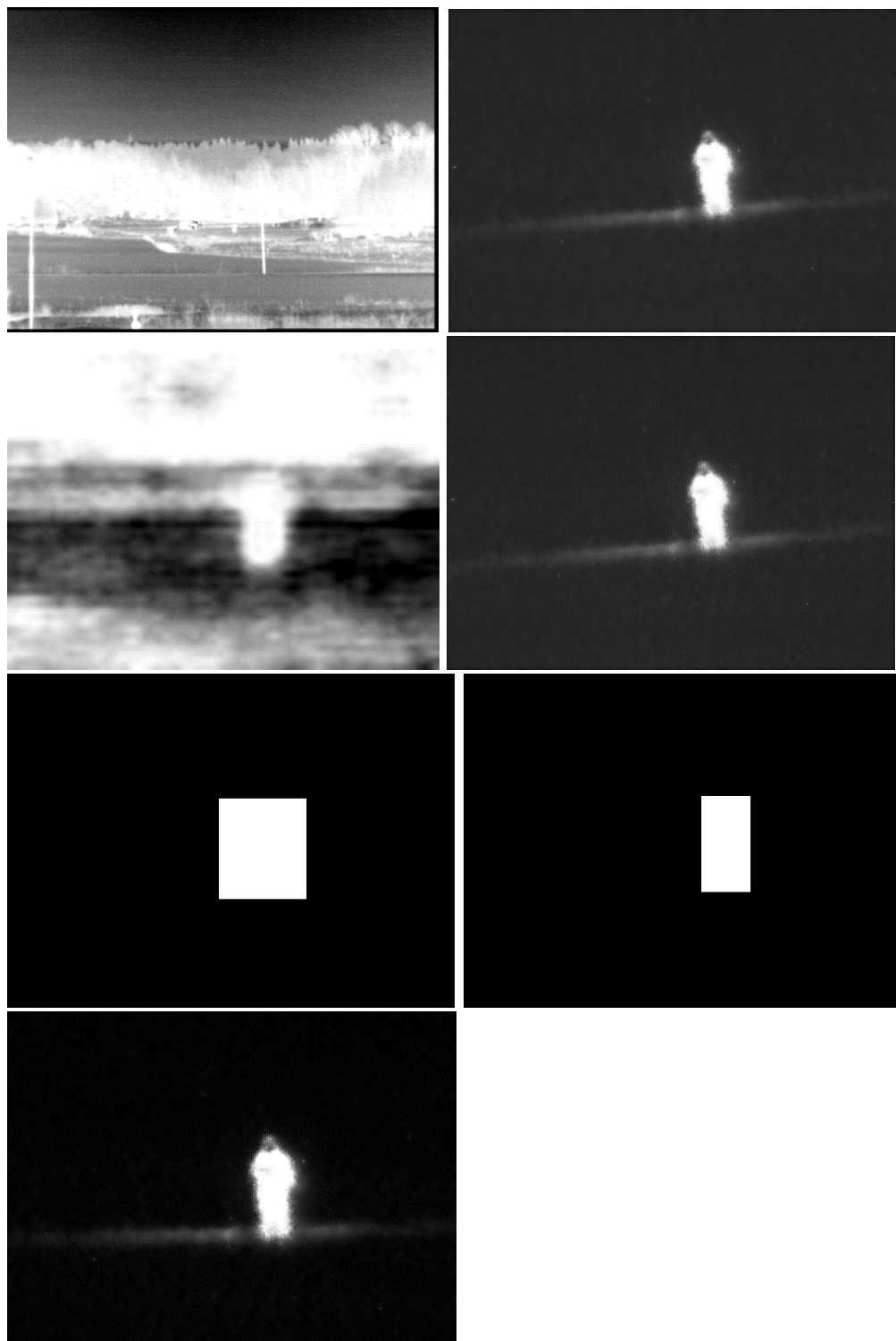
- Opt = groundtruth image optimale fit, minimale scaling
- Opt2 = groundtruth image voor meest optimale fit, meer scaling
- Best = absolute fit als groundtruth, edoch GV is erg deformed terwijl aanname is de ratio in GV niet te wijzigen
- Unscaled = unscaled variant als groundtruth (past het beste binnen aanname dat scaling niet nodig hoort te zijn indien FOV is meegenomen in registratie algoritme)

Tabel met resultaten registratie													
filename	cc opt	cc opt2	cc best	cc unscaled	mu diff opt	mu diff opt2	mu diff best	mu diff unscaled	std diff opt	std diff opt2	std diff best	std diff unscaled	
Optimalfit.tif	1,0000	0,8084	0,8101	0,5376	0,0000	4,3037	4,9020	11,7529	0,0000	18,9644	19,7713	43,1607	
Optimalfit2.tif	0,8084	1,0000	0,8974	0,4177	4,3037	0,0000	4,0778	13,9756	18,9644	0,0000	14,2619	46,5026	
Bestfit.tif	0,8101	0,8974	1,0000	0,4235	4,3020	4,0778	0,0000	14,3680	19,7713	14,2619	0,0000	46,7166	
Unscaledfit.tif	0,5376	0,4177	0,4235	1,0000	11,7529	13,9756	14,3680	0,0000	43,1607	46,5026	46,7166	0,0000	
Rigid1.tif	0,7331	0,9239	0,8341	0,3354	5,9071	4,4639	5,7208	14,2483	21,6529	10,4620	17,7976	48,7411	
Rigid2.tif	0,7270	0,9231	0,8311	0,3167	6,2778	4,8516	6,0566	14,6732	21,7506	10,2774	17,7224	49,0159	
Rigid3.tif	0,7128	0,9201	0,8336	0,3061	6,0141	4,5751	5,7712	14,2831	22,4551	10,7182	17,8290	49,6106	
Rigid4.tif	0,7128	0,9201	0,8336	0,3061	6,0141	4,5751	5,7712	14,2831	22,4551	10,7182	17,8290	49,6106	
Rigid5.tif	0,7076	0,9189	0,8303	0,2809	6,3934	5,0057	6,1376	14,7424	22,5295	10,5345	17,7516	50,0851	
Rigid6.tif	0,7204	0,9219	0,8333	0,3158	5,9489	4,5324	5,7596	14,2616	22,1628	10,5945	17,8383	49,3248	
Rigid7.tif	0,7073	0,9188	0,8299	0,2782	6,3989	5,0153	6,1492	14,7425	22,5428	10,5364	17,7676	50,1679	
Rigid8.tif	0,7032	0,9178	0,8324	0,2790	6,0677	4,7019	5,8454	14,3255	22,8615	10,8565	17,8823	50,4299	
Similarity1.tif	0,7895	0,9788	0,8937	0,4162	34,0235	32,5673	34,2608	42,5701	19,7224	6,6333	14,7009	47,3677	
Similarity2.tif	0,7792	0,9779	0,8940	0,4029	36,0534	34,3343	36,1040	44,7463	20,1304	7,0251	14,4554	47,6700	
Similarity3.tif	0,7789	0,9785	0,8948	0,4020	33,8246	32,0637	33,8571	42,5865	20,0924	6,7007	14,2567	47,6166	
Similarity4.tif	0,7877	0,9826	0,8950	0,4103	33,8590	32,2346	34,0340	42,5172	19,7818	6,3777	14,4504	47,4577	
Tabel met settings elastix													
filename	metric	iteration	hist bins	spat samples	Sp-a	Sp-alpha	Sp-A	numb resol					
Optimalfit.tif													
Optimalfit2.tif													
Bestfit.tif													
Unscaledfit.tif													
Rigid1.tif	MI	500	64	3000	1000	0,602		50				4	
Rigid2.tif	MI	1000	64	3000	1000	0,602		50				4	
Rigid3.tif	MI	2000	64	3000	1000	0,602		50				4	
Rigid4.tif	MI	2000	64	3000	10000	0,602		50				4	
Rigid5.tif	MI	3000	64	3000	1000	0,602		50				4	
Rigid6.tif	NMI	3000	64	3000	1000	0,602		50				4	
Rigid7.tif	MI	3000	64	3000	1000	0,602		50				6	
Rigid8.tif	MI	3000	64	4000	1000	0,602		50				6	
Similarity1.tif	MI	2000	64	3000	1000	0,602		50				4	
Similarity2.tif	MI	3000	64	3000	1000	0,602		50				4	
Similarity3.tif	MI	3000	64	3000	1000	0,602		50				6	
Similarity4.tif	MI	3000	64	4000	1000	0,602		50				6	

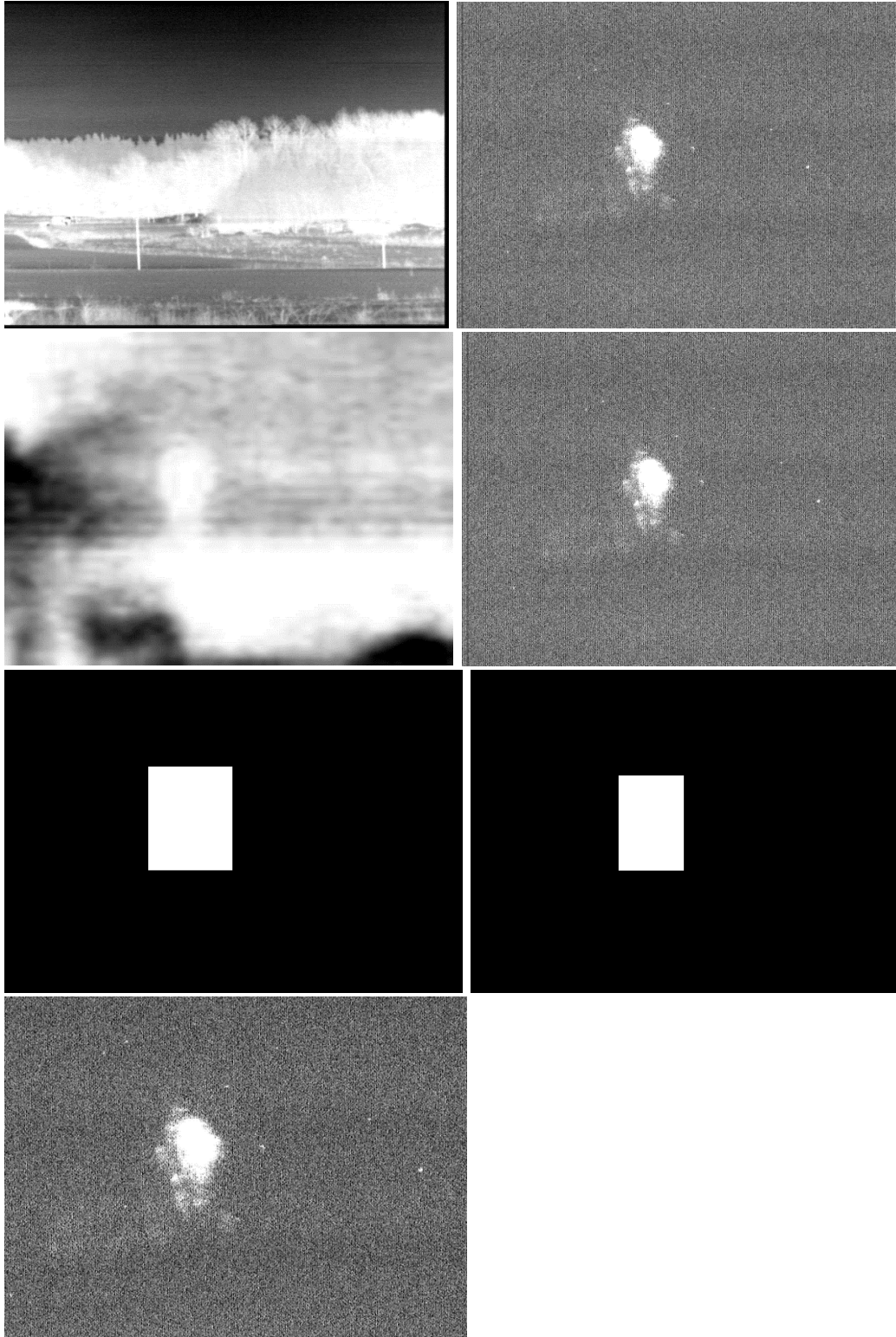
Appendix B: Resultaten registratie scenario 1 en 7

Per scenario: IR, GV, fixedImage, movingImage, fixedMask, movingMask en elastix result

Scenario 1:



Scenario 7:



Appendix C: tabellen met resultaten IQM op beelden gebruikt voor human in the loop validatie en met correlatie

IQM results selection of 10 fused images of scenario 1

Image	Fusion method	Fusion strategy	LF fusionrule	HF fusionrule	IQM W&B FTS	IQM Piella FTS
01	contourlet	pixelbased	0.5	maxabs	0.8483	0.0717
02	weighted Avg	pixelbased	0.5		0.8496	0.0968
03	weighted Avg	pixelbased	PCA		0.6830	0.0719
04	contourlet	pixelbased	locmax	maxabs	0.9447	0.0549
05	priority fused				0.9549	0.0041
06	contourlet	regionbased	locvar	maxabs	0.9545	0.0408
07	contourlet	regionbased	0.5	maxabs	0.9463	0.0362
08	weighted Avg	regionbased	locmax		0.9418	0.0798
IR	IR non fused				0.9330	0.8451

Image	Fusion method	Fusion strategy	LF fusionrule	HF fusionrule	IQM W&B locvar	IQM Piella locvar
01	contourlet	pixelbased	0.5	maxabs	0.6482	0.0548
02	weighted Avg	pixelbased	0.5		0.6459	0.0736
03	weighted Avg	pixelbased	PCA		0.7320	0.0771
04	contourlet	pixelbased	locmax	maxabs	0.6404	0.0372
05	priority fused				0.6976	0.0030
06	contourlet	regionbased	locvar	maxabs	0.6947	0.0297
07	contourlet	regionbased	0.5	maxabs	0.6458	0.0247
08	weighted Avg	regionbased	locmax		0.6196	0.0525
IR	IR non fused				0.5910	0.5353

Image	Fusion method	Fusion strategy	LF fusionrule	HF fusionrule	IQM W&B harris	IQM Piella harris
01	contourlet	pixelbased	0.5	maxabs	0.7390	0.0625
02	weighted Avg	pixelbased	0.5		0.7378	0.0841
03	weighted Avg	pixelbased	PCA		0.4256	0.0448
04	contourlet	pixelbased	locmax	maxabs	0.9350	0.0543
05	priority fused				0.9810	0.0042
06	contourlet	regionbased	locvar	maxabs	0.9799	0.0419
07	contourlet	regionbased	0.5	maxabs	0.9461	0.0362
08	weighted Avg	regionbased	locmax		0.9274	0.0786
IR	IR non fused				0.9003	0.8155

Image	Fusion method	Fusion strategy	LF fusionrule	HF fusionrule	IQM W&B harris2	IQM Piella harris2
01	contourlet	pixelbased	0.5	maxabs	0.7833	0.0662
02	weighted Avg	pixelbased	0.5		0.7799	0.0889
03	weighted Avg	pixelbased	PCA		0.9203	0.0969
04	contourlet	pixelbased	locmax	maxabs	0.7386	0.0429
05	priority fused				0.8236	0.0035
06	contourlet	regionbased	locvar	maxabs	0.8202	0.0351
07	contourlet	regionbased	0.5	maxabs	0.7542	0.0288
08	weighted Avg	regionbased	locmax		0.7187	0.0609
IR	IR non fused				0.6774	0.6136

Image	Fusion method	Fusion strategy	LF fusionrule	HF fusionrule	IQM W&B itti koch	IQM Piella itti koch
01	contourlet	pixelbased	0.5	maxabs	0.5794	0.0490
02	weighted Avg	pixelbased	0.5		0.5828	0.0664
03	weighted Avg	pixelbased	PCA		0.2093	0.0220
04	contourlet	pixelbased	locmax	maxabs	0.7994	0.0464
05	priority fused				0.7987	0.0034
06	contourlet	regionbased	locvar	maxabs	0.7988	0.0341
07	contourlet	regionbased	0.5	maxabs	0.7994	0.0306
08	weighted Avg	regionbased	locmax		0.7987	0.0677
IR	IR non fused				0.7972	0.7221

Human ranking selection of 10 fused images of scenario 1

Image:	1	2	3	4	5	6	7	8	IR
Ranking:	7.4	7.4	8.4	2.9	5.6	5.7	4.4	2.0	1.1

Correlation IQM-Human

Method:	W&B+FTS	W&B+lv	W&B+h	W&B+h2	W&B+ik	P+FTS	P+lv	P+h	P+h2	P+ik
correlation all:	-0.6943	0.7499	-0.6629	0.8437	-0.7364	-0.5480	-0.5099	-0.5701	-0.4992	-0.5846
correlation no IR:	-0.7437	0.6296	-0.7263	0.7746	-0.7672	0.1864	0.4371	-0.0711	0.4764	-0.2674
correlation pixelbased:	-0.8195	0.5680	-0.8349	0.7113	-0.8226	0.6527	0.8951	0.1247	0.9012	-0.1925
correlation regionbased:	0.9495	0.9472	0.9509	0.9480	0.0633	-0.8588	-0.8250	-0.8457	-0.8177	-0.8632
rankcorr all:	-0.4268	0.8452	-0.4268	0.8452	-0.5823	-0.1255	0.1088	-0.2510	0.1088	-0.4770
rankcorr no IR:	-0.5389	0.7785	-0.5389	0.7785	-0.7637	0.2515	0.5868	0.0719	0.5868	-0.2515
rankcorr pixelbased:	-0.9487	0.9487	-0.9487	0.9487	-0.9487	0.6325	0.9487	-0.3162	0.9487	-0.3162
rankcorr regionbased:	0.8000	0.8000	0.8000	0.8000	0.2108	-0.4000	-0.4000	-0.4000	-0.4000	-0.4000

W&B = Wang&Bovik, P = Piella, lv = local variance, h = harris, ik = itti Koch

Appendix D: resultaten IQM op alle scenarios

In de resultaten steeds naam van de fusiealgoritme gevolgd door IQM waarde. Naam: methode_patch_strategie_LF-fusieregels_HF-fusieregels. Patch betekent het eigenlijke fusie deel in het IR beeld. Bij weighted fusion HF en LF fusie regel is vervangen door weighting methode. Bij niet gefuseerd en priority fusie spreekt de naam voor zich. De beste resultaten van scenario 1 zijn vervolgens toegepast op de overige scenario's.

Scenario 1: IQM op alle fusie resultaten en selectie van de beste 8

Resultaten pixelbased: Piella met harris2 (rood selectie, geel de beste)

IQM Piella:	GV1.tif	0.1083
IQM Piella:	IR1patch.tif	0.6136
IQM Piella:	contourlet_patch_pixelbased_0.5_average.tif	0.0580
IQM Piella:	contourlet_patch_pixelbased_0.5_maxabs.tif	0.0662
IQM Piella:	contourlet_patch_pixelbased_0.5_maxaverage.tif	0.0526
IQM Piella:	contourlet_patch_pixelbased_0.5_maxlocabs.tif	0.0552
IQM Piella:	contourlet_patch_pixelbased_0.5_maxlocvar.tif	0.0443
IQM Piella:	contourlet_patch_pixelbased_locmax_average.tif	0.0463
IQM Piella:	contourlet_patch_pixelbased_locmax_maxabs.tif	0.0429
IQM Piella:	contourlet_patch_pixelbased_locmax_maxaverage.tif	0.0879
IQM Piella:	contourlet_patch_pixelbased_locmax_maxlocabs.tif	0.0273
IQM Piella:	contourlet_patch_pixelbased_locmax_maxlocvar.tif	0.0799
IQM Piella:	curvelet_patch_pixelbased_0.5_average.tif	0.0491
IQM Piella:	curvelet_patch_pixelbased_0.5_maxabs.tif	0.0707
IQM Piella:	curvelet_patch_pixelbased_0.5_maxaverage.tif	0.0598
IQM Piella:	curvelet_patch_pixelbased_0.5_maxlocabs.tif	0.1075
IQM Piella:	curvelet_patch_pixelbased_0.5_maxlocvar.tif	0.0602
IQM Piella:	curvelet_patch_pixelbased_locmax_average.tif	0.0752
IQM Piella:	curvelet_patch_pixelbased_locmax_maxabs.tif	-0.0486
IQM Piella:	curvelet_patch_pixelbased_locmax_maxaverage.tif	0.0663
IQM Piella:	curvelet_patch_pixelbased_locmax_maxlocabs.tif	0.0138
IQM Piella:	curvelet_patch_pixelbased_locmax_maxlocvar.tif	0.0088
IQM Piella:	dtcwt_patch_pixelbased_0.5_average.tif	0.0571
IQM Piella:	dtcwt_patch_pixelbased_0.5_maxabs.tif	0.0616
IQM Piella:	dtcwt_patch_pixelbased_0.5_maxaverage.tif	0.0593
IQM Piella:	dtcwt_patch_pixelbased_0.5_maxlocabs.tif	0.0291
IQM Piella:	dtcwt_patch_pixelbased_0.5_maxlocvar.tif	0.0593
IQM Piella:	dtcwt_patch_pixelbased_locmax_average.tif	0.0630
IQM Piella:	dtcwt_patch_pixelbased_locmax_maxabs.tif	0.0691
IQM Piella:	dtcwt_patch_pixelbased_locmax_maxaverage.tif	0.0866
IQM Piella:	dtcwt_patch_pixelbased_locmax_maxlocabs.tif	0.0771
IQM Piella:	dtcwt_patch_pixelbased_locmax_maxlocvar.tif	0.0594
IQM Piella:	weightedAvg_patch_pixelbased_0.5.tif	0.0889
IQM Piella:	weightedAvg_patch_pixelbased_PCA.tif	0.0969
IQM Piella:	weightedAvg_patch_pixelbased_locmax.tif	0.1028

Resultaten regionbased: w&b met harris (rood selectie, geel de beste, noot: gezien meerdere gelijke waardes uiteindelijk 10)

IQM W&B:	GV1.tif	0.3955	IQM
IQM W&B:	IR1patch.tif	0.9003	IQM
IQM W&B:	contourlet_patch_regionbased_0.5_average.tif	0.9436	IQM
IQM W&B:	contourlet_patch_regionbased_0.5_maxabs.tif	0.9461	IQM
IQM W&B:	contourlet_patch_regionbased_0.5_maxaverage.tif	0.9448	IQM
IQM W&B:	contourlet_patch_regionbased_0.5_maxlocabs.tif	0.9461	IQM
IQM W&B:	contourlet_patch_regionbased_0.5_maxlocvar.tif	0.9461	IQM
IQM W&B:	contourlet_patch_regionbased_locmax_average.tif	0.9349	IQM
IQM W&B:	contourlet_patch_regionbased_locmax_maxabs.tif	0.9375	IQM
IQM W&B:	contourlet_patch_regionbased_locmax_maxaverage.tif	0.9363	IQM
IQM W&B:	contourlet_patch_regionbased_locmax_maxlocabs.tif	0.9374	IQM
IQM W&B:	contourlet_patch_regionbased_locmax_maxlocvar.tif	0.9374	IQM
IQM W&B:	contourlet_patch_regionbased_locvar_average.tif	0.9771	IQM
IQM W&B:	contourlet_patch_regionbased_locvar_maxabs.tif	0.9799	IQM
IQM W&B:	contourlet_patch_regionbased_locvar_maxaverage.tif	0.9783	IQM
IQM W&B:	contourlet_patch_regionbased_locvar_maxlocabs.tif	0.9799	IQM
IQM W&B:	contourlet_patch_regionbased_locvar_maxlocvar.tif	0.9799	IQM
IQM W&B:	contourlet_patch_regionbased_maxlocvar_average.tif	0.9776	IQM

IQM W&B: contourlet_patch_regionbased_maxlocvar_maxabs.tif	0.9804	IQM
IQM W&B: contourlet_patch_regionbased_maxlocvar_maxaverage.tif	0.9787	IQM
IQM W&B: contourlet_patch_regionbased_maxlocvar_maxlocabs.tif	0.9804	IQM
IQM W&B: contourlet_patch_regionbased_maxlocvar_maxlocvar.tif	0.9804	IQM
IQM W&B: curvelet_patch_regionbased_0.5_average.tif	0.9438	IQM
IQM W&B: curvelet_patch_regionbased_0.5_maxabs.tif	0.9452	IQM
IQM W&B: curvelet_patch_regionbased_0.5_maxaverage.tif	0.9438	IQM
IQM W&B: curvelet_patch_regionbased_0.5_maxlocabs.tif	0.9452	IQM
IQM W&B: curvelet_patch_regionbased_0.5_maxlocvar.tif	0.9452	IQM
IQM W&B: curvelet_patch_regionbased_locmax_average.tif	0.9303	IQM
IQM W&B: curvelet_patch_regionbased_locmax_maxabs.tif	0.9316	IQM
IQM W&B: curvelet_patch_regionbased_locmax_maxaverage.tif	0.9303	IQM
IQM W&B: curvelet_patch_regionbased_locmax_maxlocabs.tif	0.9316	IQM
IQM W&B: curvelet_patch_regionbased_locmax_maxlocvar.tif	0.9316	IQM
IQM W&B: curvelet_patch_regionbased_locvar_average.tif	0.9794	IQM
IQM W&B: curvelet_patch_regionbased_locvar_maxabs.tif	0.9806	IQM
IQM W&B: curvelet_patch_regionbased_locvar_maxaverage.tif	0.9792	IQM
IQM W&B: curvelet_patch_regionbased_locvar_maxlocabs.tif	0.9806	IQM
IQM W&B: curvelet_patch_regionbased_locvar_maxlocvar.tif	0.9806	IQM
IQM W&B: curvelet_patch_regionbased_maxlocvar_average.tif	0.9799	IQM
IQM W&B: curvelet_patch_regionbased_maxlocvar_maxabs.tif	0.9810	IQM
IQM W&B: curvelet_patch_regionbased_maxlocvar_maxaverage.tif	0.9797	IQM
IQM W&B: curvelet_patch_regionbased_maxlocvar_maxlocabs.tif	0.9810	IQM
IQM W&B: curvelet_patch_regionbased_maxlocvar_maxlocvar.tif	0.9810	IQM
IQM W&B: dtcwt_patch_regionbased_0.5_average.tif	0.9436	IQM
IQM W&B: dtcwt_patch_regionbased_0.5_maxabs.tif	0.9471	IQM
IQM W&B: dtcwt_patch_regionbased_0.5_maxaverage.tif	0.9432	IQM
IQM W&B: dtcwt_patch_regionbased_0.5_maxlocabs.tif	0.9471	IQM
IQM W&B: dtcwt_patch_regionbased_0.5_maxlocvar.tif	0.9471	IQM
IQM W&B: dtcwt_patch_regionbased_locmax_average.tif	0.9326	IQM
IQM W&B: dtcwt_patch_regionbased_locmax_maxabs.tif	0.9362	IQM
IQM W&B: dtcwt_patch_regionbased_locmax_maxaverage.tif	0.9322	IQM
IQM W&B: dtcwt_patch_regionbased_locmax_maxlocabs.tif	0.9362	IQM
IQM W&B: dtcwt_patch_regionbased_locmax_maxlocvar.tif	0.9362	IQM
IQM W&B: dtcwt_patch_regionbased_locvar_average.tif	0.9758	IQM
IQM W&B: dtcwt_patch_regionbased_locvar_maxabs.tif	0.9789	IQM
IQM W&B: dtcwt_patch_regionbased_locvar_maxaverage.tif	0.9750	IQM
IQM W&B: dtcwt_patch_regionbased_locvar_maxlocabs.tif	0.9789	IQM
IQM W&B: dtcwt_patch_regionbased_locvar_maxlocvar.tif	0.9789	IQM
IQM W&B: dtcwt_patch_regionbased_maxlocvar_average.tif	0.9762	IQM
IQM W&B: dtcwt_patch_regionbased_maxlocvar_maxabs.tif	0.9793	IQM
IQM W&B: dtcwt_patch_regionbased_maxlocvar_maxaverage.tif	0.9755	IQM
IQM W&B: dtcwt_patch_regionbased_maxlocvar_maxlocabs.tif	0.9793	IQM
IQM W&B: dtcwt_patch_regionbased_maxlocvar_maxlocvar.tif	0.9793	IQM
IQM W&B: priorityfused_patch.tif	0.9810	IQM
IQM W&B: weightedAvg_patch_regionbased_0.5.tif	0.9437	IQM
IQM W&B: weightedAvg_patch_regionbased_PCA.tif	0.9784	IQM
IQM W&B: weightedAvg_patch_regionbased_locmax.tif	0.9274	IQM
IQM W&B: weightedAvg_patch_regionbased_locvar.tif	0.9797	IQM

Scenario 2

Resultaten pixelbased

IQM Piella: contourlet_patch_pixelbased_locmax_maxaverage.tif	0.1293
IQM Piella: contourlet_patch_pixelbased_locmax_maxlocvar.tif	0.0187
IQM Piella: curvelet_patch_pixelbased_0.5_maxlocabs.tif	0.0342
IQM Piella: dtcwt_patch_pixelbased_locmax_maxaverage.tif	0.0212
IQM Piella: dtcwt_patch_pixelbased_locmax_maxlocabs.tif	0.0181
IQM Piella: weightedAvg_patch_pixelbased_0.5.tif	0.0696
IQM Piella: weightedAvg_patch_pixelbased_PCA.tif	0.0647
IQM Piella: weightedAvg_patch_pixelbased_locmax.tif	0.0699

Resultaten regionbased

IQM W&B: contourlet_patch_regionbased_maxlocvar_maxabs.tif	0.5152
IQM W&B: contourlet_patch_regionbased_maxlocvar_maxlocabs.tif	0.5152
IQM W&B: contourlet_patch_regionbased_maxlocvar_maxlocvar.tif	0.5152
IQM W&B: curvelet_patch_regionbased_locvar_maxabs.tif	0.5152
IQM W&B: curvelet_patch_regionbased_locvar_maxlocabs.tif	0.5152

IQM W&B:	curvelet_patch_regionbased_locvar_maxlocvar.tif	0.5152
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxabs.tif	0.5153
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocabs.tif	0.5153
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocvar.tif	0.5153
IQM W&B:	priorityfused_patch_.tif	0.5153

Scenario 3

Resultaten pixelbased

IQM Piella:	contourlet_patch_pixelbased_locmax_maxaverage.tif	0.0433
IQM Piella:	contourlet_patch_pixelbased_locmax_maxlocvar.tif	-0.0149
IQM Piella:	curvelet_patch_pixelbased_0.5_maxlocabs.tif	-0.0201
IQM Piella:	dtcwt_patch_pixelbased_locmax_maxaverage.tif	0.0135
IQM Piella:	dtcwt_patch_pixelbased_locmax_maxlocabs.tif	0.3336
IQM Piella:	weightedAvg_patch_pixelbased_0.5.tif	0.1479
IQM Piella:	weightedAvg_patch_pixelbased_PCA.tif	0.1400
IQM Piella:	weightedAvg_patch_pixelbased_locmax.tif	0.1062

Resultaten regionbased

IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxabs.tif	0.7460
IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxlocabs.tif	0.7460
IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxlocvar.tif	0.7460
IQM W&B:	curvelet_patch_regionbased_locvar_maxabs.tif	0.7458
IQM W&B:	curvelet_patch_regionbased_locvar_maxlocabs.tif	0.7458
IQM W&B:	curvelet_patch_regionbased_locvar_maxlocvar.tif	0.7458
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxabs.tif	0.7460
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocabs.tif	0.7460
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocvar.tif	0.7460
IQM W&B:	priorityfused_patch_.tif	0.7460

Scenario 4

Resultaten pixelbased

IQM Piella:	contourlet_patch_pixelbased_locmax_maxaverage.tif	-0.0065
IQM Piella:	contourlet_patch_pixelbased_locmax_maxlocvar.tif	0.0098
IQM Piella:	curvelet_patch_pixelbased_0.5_maxlocabs.tif	0.0755
IQM Piella:	dtcwt_patch_pixelbased_locmax_maxaverage.tif	0.0649
IQM Piella:	dtcwt_patch_pixelbased_locmax_maxlocabs.tif	0.1345
IQM Piella:	weightedAvg_patch_pixelbased_0.5.tif	0.1650
IQM Piella:	weightedAvg_patch_pixelbased_PCA.tif	0.2961
IQM Piella:	weightedAvg_patch_pixelbased_locmax.tif	0.0623

Resultaten regionbased

IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxabs.tif	0.6038
IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxlocabs.tif	0.6038
IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxlocvar.tif	0.6038
IQM W&B:	curvelet_patch_regionbased_locvar_maxabs.tif	0.6038
IQM W&B:	curvelet_patch_regionbased_locvar_maxlocabs.tif	0.6038
IQM W&B:	curvelet_patch_regionbased_locvar_maxlocvar.tif	0.6038
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxabs.tif	0.6038
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocabs.tif	0.6038
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocvar.tif	0.6038
IQM W&B:	priorityfused_patch_.tif	0.6038

Scenario 5

Resultaten pixelbased

IQM Piella:	contourlet_patch_pixelbased_locmax_maxaverage.tif	0.0184
IQM Piella:	contourlet_patch_pixelbased_locmax_maxlocvar.tif	-0.0020
IQM Piella:	curvelet_patch_pixelbased_0.5_maxlocabs.tif	-0.0090
IQM Piella:	dtcwt_patch_pixelbased_locmax_maxaverage.tif	0.1826
IQM Piella:	dtcwt_patch_pixelbased_locmax_maxlocabs.tif	0.3300
IQM Piella:	weightedAvg_patch_pixelbased_0.5.tif	0.1408
IQM Piella:	weightedAvg_patch_pixelbased_PCA.tif	0.1080
IQM Piella:	weightedAvg_patch_pixelbased_locmax.tif	-0.0392

Resultaten regionbased

IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxabs.tif	0.4398
IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxlocabs.tif	0.4398
IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxlocvar.tif	0.4398

IQM W&B:	curvelet_patch_regionbased_locvar_maxabs.tif	0.4395
IQM W&B:	curvelet_patch_regionbased_locvar_maxlocabs.tif	0.4395
IQM W&B:	curvelet_patch_regionbased_locvar_maxlocvar.tif	0.4395
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxabs.tif	0.4400
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocabs.tif	0.4400
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocvar.tif	0.4400
IQM W&B:	priorityfused_patch_.tif	0.4400

Scenario 6

Resultaten pixelbased

IQM Piella:	contourlet_patch_pixelbased_locmax_maxaverage.tif	0.0151
IQM Piella:	contourlet_patch_pixelbased_locmax_maxlocvar.tif	0.0002
IQM Piella:	curvelet_patch_pixelbased_0.5_maxlocabs.tif	0.6243
IQM Piella:	dctwt_patch_pixelbased_locmax_maxaverage.tif	0.4685
IQM Piella:	dctwt_patch_pixelbased_locmax_maxlocabs.tif	0.4145
IQM Piella:	weightedAvg_patch_pixelbased_0.5.tif	0.0292
IQM Piella:	weightedAvg_patch_pixelbased_PCA.tif	0.2297
IQM Piella:	weightedAvg_patch_pixelbased_locmax.tif	0.1299

Resultaten regionbased

IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxabs.tif	0.6992
IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxlocabs.tif	0.6992
IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxlocvar.tif	0.6992
IQM W&B:	curvelet_patch_regionbased_locvar_maxabs.tif	0.6991
IQM W&B:	curvelet_patch_regionbased_locvar_maxlocabs.tif	0.6991
IQM W&B:	curvelet_patch_regionbased_locvar_maxlocvar.tif	0.6991
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxabs.tif	0.6993
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocabs.tif	0.6993
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocvar.tif	0.6993
IQM W&B:	priorityfused_patch_.tif	0.6993

Scenario 7

Resultaten pixelbased

IQM Piella:	contourlet_patch_pixelbased_locmax_maxaverage.tif	-0.1895
IQM Piella:	contourlet_patch_pixelbased_locmax_maxlocvar.tif	0.0342
IQM Piella:	curvelet_patch_pixelbased_0.5_maxlocabs.tif	0.1524
IQM Piella:	dctwt_patch_pixelbased_locmax_maxaverage.tif	-0.1656
IQM Piella:	dctwt_patch_pixelbased_locmax_maxlocabs.tif	0.1288
IQM Piella:	weightedAvg_patch_pixelbased_0.5.tif	0.1201
IQM Piella:	weightedAvg_patch_pixelbased_PCA.tif	0.2483
IQM Piella:	weightedAvg_patch_pixelbased_locmax.tif	0.0193

Resultaten regionbased

IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxabs.tif	0.5964
IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxlocabs.tif	0.5964
IQM W&B:	contourlet_patch_regionbased_maxlocvar_maxlocvar.tif	0.5964
IQM W&B:	curvelet_patch_regionbased_locvar_maxabs.tif	0.5963
IQM W&B:	curvelet_patch_regionbased_locvar_maxlocabs.tif	0.5963
IQM W&B:	curvelet_patch_regionbased_locvar_maxlocvar.tif	0.5963
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxabs.tif	0.5964
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocabs.tif	0.5964
IQM W&B:	curvelet_patch_regionbased_maxlocvar_maxlocvar.tif	0.5964
IQM W&B:	priorityfused_patch_.tif	0.5964

Ranking IQM waarden

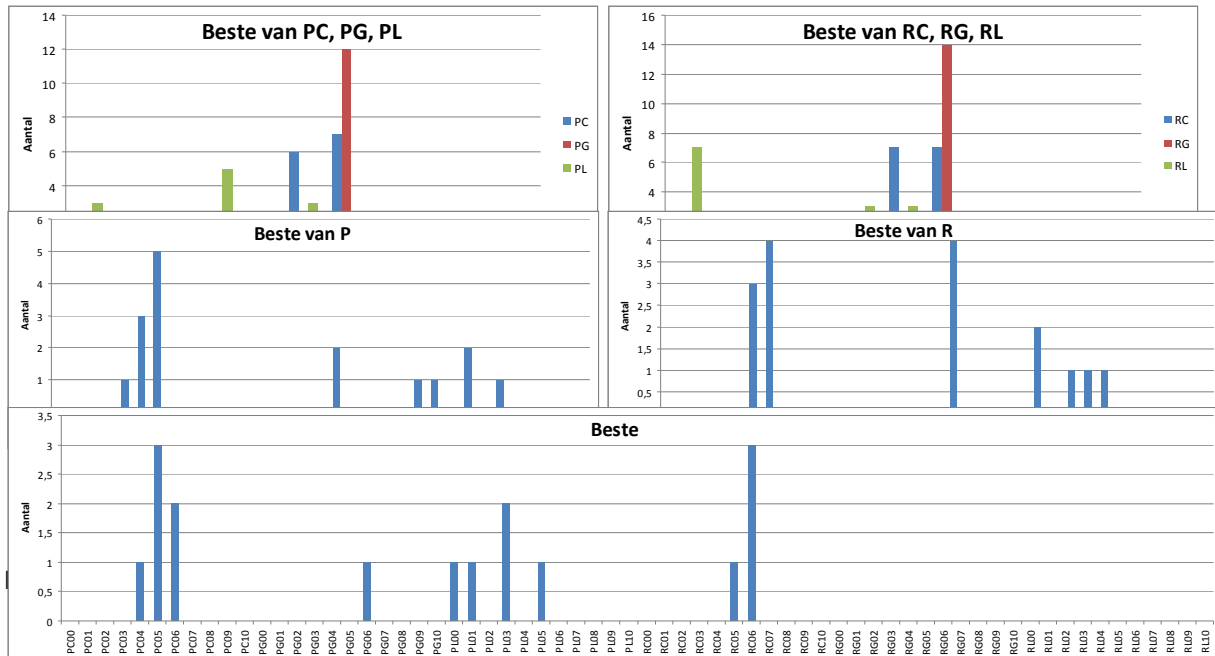
Geel gearceerd steeds de beste.

Ranking pixelbased										
scenarios	1	2	3	4	5	6	7	average rank	rank	
contourlet_patch_pixelbased_locmax_maxaverage.tif	5	1	5	8	5	7	8	5,57	7	
contourlet_patch_pixelbased_locmax_maxlocvar.tif	7	7	7	7	6	8	5	6,71	8	
curvelet_patch_pixelbased_0.5_maxlocabs.tif	1	5	8	4	7	1	2	4,00	4	
dtcwt_patch_pixelbased_locmax_maxaverage.tif	6	6	6	5	2	2	7	4,86	6	
dtcwt_patch_pixelbased_locmax_maxlocabs.tif	8	8	1	3	1	3	3	3,86	3	
weightedAvg_patch_pixelbased_0.5.tif	4	3	2	2	3	6	4	3,43	2	
weightedAvg_patch_pixelbased_PCA.tif	3	4	3	1	4	4	1	2,86	1	
weightedAvg_patch_pixelbased_locmax.tif	2	2	4	6	8	5	6	4,71	5	

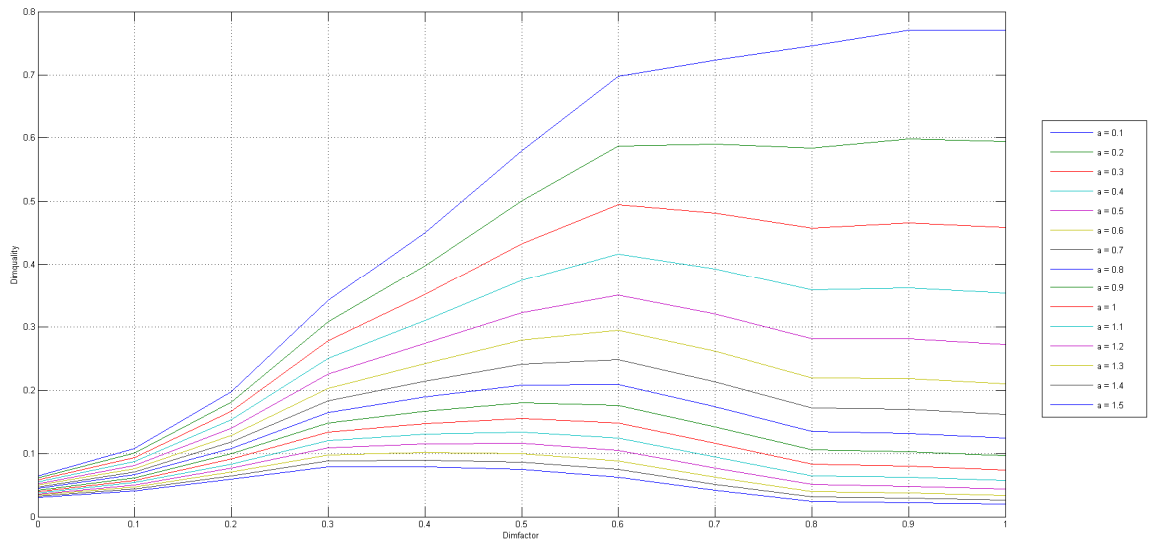
Ranking regionbased										
scenarios	1	2	3	4	5	6	7	average rank	rank	
contourlet_patch_regionbased_maxlocvar_maxabs.tif	8	5	1	1	5	5	1	3,71	5	
contourlet_patch_regionbased_maxlocvar_maxlocabs.tif	8	5	1	1	5	5	1	3,71	5	
contourlet_patch_regionbased_maxlocvar_maxlocvar.tif	8	5	1	1	5	5	1	3,71	5	
curvelet_patch_regionbased_locvar_maxabs.tif	5	5	8	1	8	8	8	6,14	8	
curvelet_patch_regionbased_locvar_maxlocabs.tif	5	5	8	1	8	8	8	6,14	8	
curvelet_patch_regionbased_locvar_maxlocvar.tif	5	5	8	1	8	8	8	6,14	8	
curvelet_patch_regionbased_maxlocvar_maxabs.tif	1	1	1	1	1	1	1	1,00	1	
curvelet_patch_regionbased_maxlocvar_maxlocabs.tif	1	1	1	1	1	1	1	1,00	1	
curvelet_patch_regionbased_maxlocvar_maxlocvar.tif	1	1	1	1	1	1	1	1,00	1	
priorityfused_patch.tif	1	1	1	1	1	1	1	1,00	1	

Appendix E: dimming en IQM dimfit

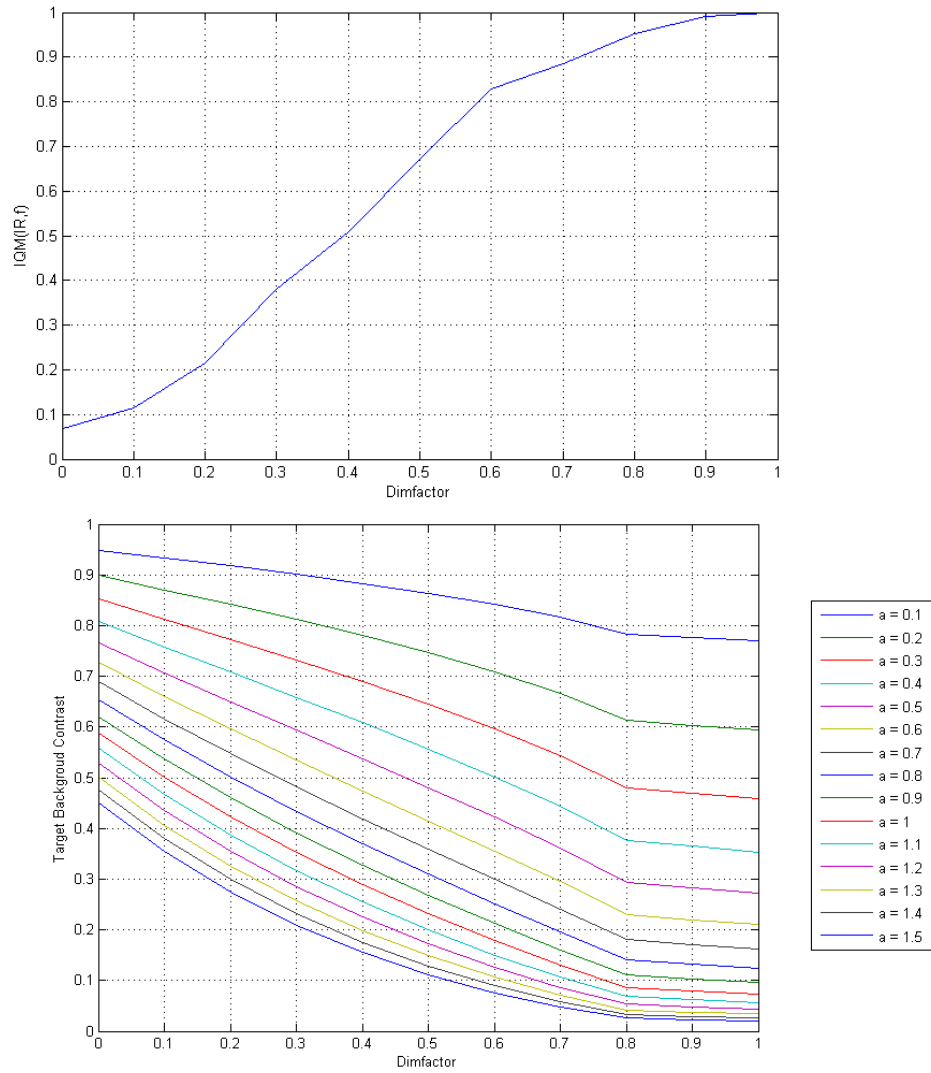
Resultaten dimming experiment



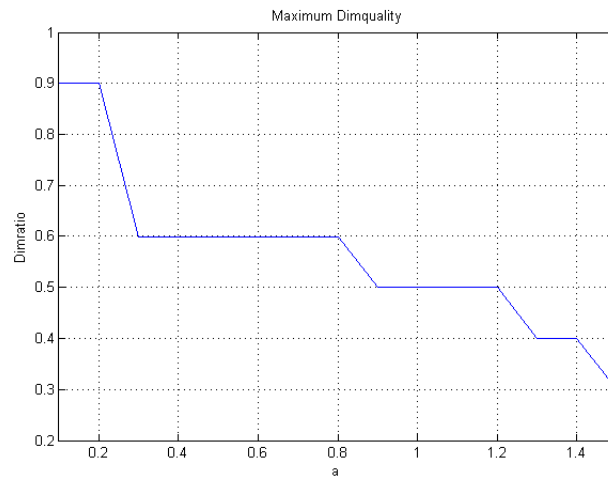
Voor steeringparametersettings a van 0.1 tot 1.5 de plots van de IQM weergegeven. Verschuivend maximum duidelijk waarneembaar.



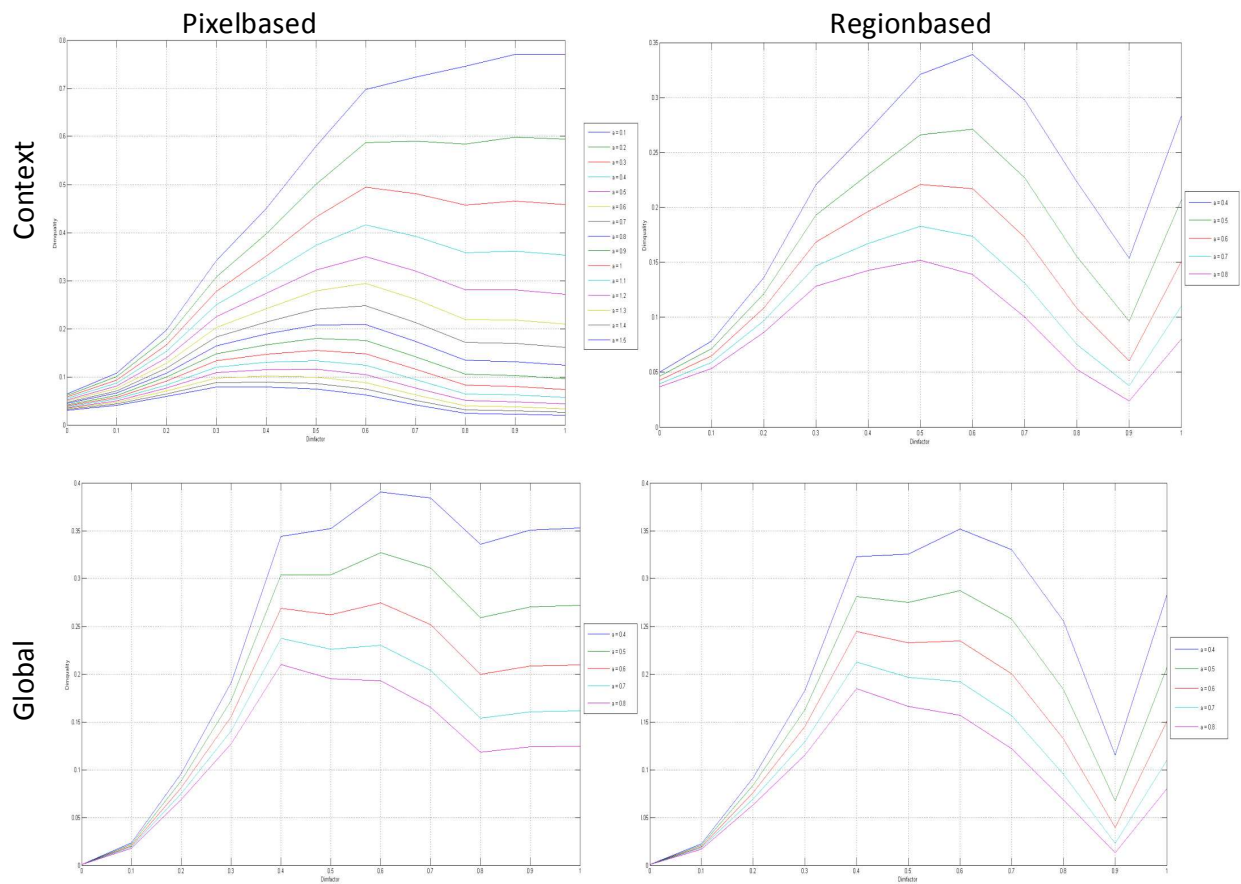
Hieronder volgt opsplitsing in W&B(IR,F) en TGT-BGR contrast.



Hieronder verschuivend maximum in IQM (dimquality) t.o.v. steeringparameter a .



Resultaten IQM op context en global dimming



Appendix F: Overview Matlab files

Type	Filename	Function
Created in the study		
Pre-processing		
	IR-preprocessing	Transfer raw data in 8-bit images
	GV-preprocessing	Apply image adjustment and de-noising (optional) with Wienerfilter
Registration		
	Registration preparation	Create input for elastix: fixed image, moving image, fixed mask, moving mask
	Registration result	Transfer elastix output into 8-bit images
	Best registration method	Apply statistics on elastix results and with a ground truth in order to define best method
Fusion		
	fusion algorithm	file to run all possible fusion algorithms, both region and pixel based
	priority fusion	algorithm for priority fusion
	weighted average fusion	algorithm for weighted fusion including fusionrules except PCA
	eigen	applying PCA
	DT-CWT	algorithm to run the DT-CWT transform module, apply fusion rules and run inverse transform
	Curvelet	algorithm to run the curvelet transform module, apply fusion rules and run inverse transform
	Contourlet	algorithm to run the contourlet transform module, apply fusion rules and run inverse transform
	GV segmentation	function to segment the human in the GV image, required for region based fusion
IQM		
	Piella & Heijmans IQM	algorithm that calculates the IQM, based on W&B
	Saliency	function to run saliency (including algorithms for variance)
Dimming		
	Dimming	Algorithm applying all three dimming methods
	Dimfit	Algorithm that calculates the dim quality for several tuning parameter settings
Video		
	makevideo	function that performs fusion of all frames, makes dimmed frames and generates the movie
Re-used files (not created)		
	DT-CWT transform	DT-CWT transform
	DT-CWT inverse transform	DT-CWT inverse transform
	curvelet transform	curvelet transform
	curvelet inverse transform	curvelet inverse transform
	contourlet transform	contourlet transform
	contourlet inverse transform	contourlet inverse transform
	Wang & Bovik IQI	Image Quality Index by Wang & Bovik (by Zhou Wang)
	FTS	Algorithm to compute frequency tuned saliency
	Harris 1	Algorithm to calculate harriscorner map using 2nd order derivative
	Harris 2	Algorithm to calculate harriscorner map using fft2
	ittikoch	Itti Koch & Niebur saliency map toolbox