They could live - encoding invisible information in displays and ways to extract it

Vojta Crha
Supervisor(s): Marco Zuñiga Zamalloa, Miguel Chavez Tapia
EEMCS, Delft University of Technology, The Netherlands

A Dissertation Submitted to EEMCS faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 19, 2022

Abstract

In the recent years, several papers have implemented methods which seamlessly integrate data streams into high frame rate monitors without affecting casual viewer experience. This paper proposes an improvement to existing systems which can in theory increase the data throughput of the data stream several times. Exploiting the same characteristics of human vision as related papers, grid containing different colours is encoded into a video. Testing was conducted to determine the possible uses of this approach. It was determined that 2x or even higher increase is possible however there seems to always be a visual artifact present. The testing also revealed a need for further testing to determine the best possible configuration (amount of colours) to use as well. This testing would also increase the reliability of the data gathered as the experiment was limited by time.

1 Introduction

Camera-to-screen is a perspective concept that can bring great utility to every phone user. On the go users can receive personalized information from virtually any screen. They can quickly share information between devices or download whole files just by pointing their phone camera at a screen. This technology can do virtually anything that can be achieved with QR codes but with several key benefits. The digital nature allows it change dynamically, either in regular intervals (for example with changing ad panels) or more rapidly, transmitting a data stream of sorts. Another key benefit is also the hidden nature of this technology. The reader can imagine it as rapidly flashing QR codes. If these QR codes are put together in just the right way and flashed fast enough, they become invisible to the human eye. A camera with sufficient capture frequency can however record these QR codes and use the information. Having such a hidden technology can have great benefits - the information is available only for people who require it and doesn't obstruct the view of others. There exists a number of implementations of this technology, such as solutions proposed in [1] , [5], [9] or [2]. These implementations however, use only one, what this paper calls, luminance value. This paper then instead proposes usage of multiple luminance values. Coming back to the QR code example, reader can imagine this as a QR code using multiple colours to encode data not as binary, but possibly ternary or even quaternary (or more realistically, sending two bits at once). This idea somewhat corresponds to an already proposed (albeit unused) QR code standard - HCC2D [10]. In theory, this approach should allow to transmit several times more data.

The paper will be structured in the following way. Section two explains the methodology used. Third section explains the contribution of this paper, as well as explains the proposed implementation in detail. Section four goes into detail about the specifics of the experiment - the way it was designed, how bias was prevented and logs how it was carried out. Fifth section handles all the ethical aspects of the research - as well as potential issues with a previously planned user research. Section six discusses the results obtained and possible uses of the technology based on the results. Last section, section seven, concludes the paper and presents possible future work that can be carried out to further explore and improve this technology.

2 Methodology

In order to fully analyze the possible benefits of using multiple luminance levels during screen-to-camera communication, a rigid approach is needed. As the goal is to maximize data throughput with little to none erceptible artifacts, the process also has to include user testing to determine whether the algorithms have any human-visible effects. Therefore, the following process was outlined - firstly, the focus will be on development of the algorithm that can determine a tuple of colours which will add up into the desired colour (desired colour being the input of the algorithm). Afterwards the focus shift to applying different luminance levels to these "composing colours" and figuring out the possible ranges of luminance that can be used without interfering with the human user experience. Lastly, this algorithm will be improved to be able to encode images. During this step, it is quite likely that more information about the varying luminance levels will be discovered as some artifacts may arise when rapidly changing between colours. Therefore further bugfixing of that part of the pipeline will be done at the same time. Lastly, the finalised algorithm will be tested by the author to determine the amount of artifacts observed between different algorithm configurations.

This testing will be conducted by the author. To ensure this data is usable, a graphical aptitude test will be taken. This will later allow to reference test results to other possible future testing participants. The specifics of the test will be described in 4.4. Originally, a larger scale test was planned but was ultimate not carried out due to reasons specified in 5.

Next, all the results will be analyzed to determine how many possible luminance levels can be used without interfering with the user experience. As this may vary across different colours due to the way human vision works (as examined in [3]), the experimental image will contain all colours to make sure the results are applicable to all real-world scenarios.

In order to properly generate and analyse the results, several tables will be created. First group part of these tables will each contain the values used for experimentation. Afterwards, a single table combining these values will be made. Following, table explaining one the scoring criteria is created and lastly, two tables presenting the results are present. The first reporting the results of experimentation with the 144Hz screen (mentioned in 4 and the other reporting about the 60Hz screen results. All of these tables can be found in section 4

3 Luminance level variation

This section will briefly explain the inner workings of existing encoding algorithms, detail the ideas for improvement which are the basis of this paper and lastly go into detail about all the implementation this paper uses as well as all related aspects.

3.1 Existing encoding algorithms

Simply said - the existing systems work by rapidly flashing slightly altered frames in rapid succession. These differences between frames are not distinguishable by humans but can be detected and used by camera devices. Now to actually understand how this is useful and what this paper wants to achieve, it is necessary to look at this bit more in depth.

While the best understanding of the existing algorithms can be acquired by reading the papers written by the authors of these algorithms, this section will provide a high-level understanding to fully grasp the improvements proposed in the following sections.

The first step is a generation of a framemask. Figure 1 - this is a grid of squares of two shades of gray. One can think of this grid as a QR code [1] with certain transparency. Afterwards a complement framemask is created - figure 2. This framemask is the exact opposite of the first mask. The next step is applying the framemasks - this is done by standard Alpha Composition [6]. After this there are three usable frames - the original unmodified, and the two affected by complementing framemasks. These frames are then repeatedly played at a frequency of at least 120 fps. At this frequency, the transition between the frames is so rapid that the human eye cannot notice the slight variation created by the framemasks. The complementing frames just "blend" into single one in the human vision. In the end, an encoded video looks identical to the source image (frame) however can be scanned by a camera and data can be extracted from the altered frames. This of course assumes perfect conditions and in practice some problems may occur. This is touched upon in 4. [2]

3.2 Ideas for improvement

As presented in the previous section, current systems rely on only one luminance value. The goal of this paper is to explore the possibilities of using multiple luminance levels. Coming back to the previously mentioned parallels with QR codes, this would mean having a QR code use multiple different colours (similar to the proposed HCC2D standard - [10]). Firstly, several aspects will be outlined. Afterwards, it will be explained how these aspects were represented in the testing and finally, the best resulting configurations will be detailed.

[1]Throughout this paper, lot of similarities with QR codes are explored. Overall, it helps to think about the proposed approach as rapidly flashing colourful QR codes

[2]The data encoding and decoding aspect is not a subject of this paper however the decoding aspect will be touched briefly as the data needs to be able to be extracted even if this paper does not attempt it. What this means in practice is that the framemasks are captured in high enough quality for the grid and the squares to be distinguishable.

Figure 1: A simple generated framemask containing 8 different colours

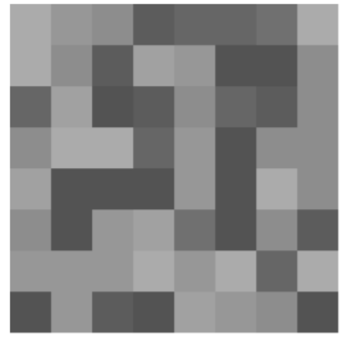
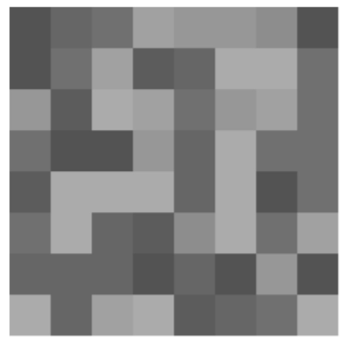


Figure 2: A complementary framemask



With this approach, there are three main aspects each of which is partially opposite to the others. All of these are related to the range of colours.

The main aspects

The first main aspect is human observability - to make sure that the encoded image looks the same as the original to the human eye. For this, ideal encoding would use a small colour range as close to the original as possible.

The second main aspect is decoding reliability. To ensure reliable data transmission, the difference between luminance values should be as big as possible and the "closest" luminance value should be as dissimilar to the original as possible. This results in some minimal distance from the original, as well as ideally a big range/small amount of luminance values.

The last main aspect is data throughput. For this, as many as possible luminance values are desired. Data throughput itself is not directly affected by other factors however it is indirectly linked to reliability and therefore the range should ideally be as wide as possible to get the most out of data transfer.

Combining all of these aspects, the range of colours should be as wide as possible but stopping at the point of human observability. There should also be as many values as possible but with differences big enough to be reliably distinguished by a camera. There will also be some initial gap between the closest value and the original. There is obviously a lot of unknowns in here - number of values, colour range, width etc. The ideal combination of these values will be determined in the results section 4.7.

3.3 Implementation details

This section will give a rough overview of the algorithm. For implementation details, the reader can use the public github repository [4] as well as the code documentation. In theory, the code is quite simple. It follows majority of principles outlined in 3.1 with an important exception to the number of colours present in the grid. That is however quite theoretical and there are many possible ways to implement such an approach. Following subsections handle the specific details of the implementation.

Overlay mask generation

As mentioned, the first step is generation of an overlay mask. First, it is important to know the dimensions of the original image. Using this information, a smaller corresponding two dimensional array is created where each entry corresponds to one square on the final grid (in the testing, a 256x256 image was used where each square corresponded to 32x32 segment). This array is then filled with the colours from the corresponding range. In the current application this is done randomly however in a real world application, this would correspond to an encoding of information. Afterwards, an array with the size of the original is filled corresponding to the previously established mapping - NxN segments filled with colours from the range.

Complementing mask generation

This step is quite simple assuming the existence of the overlay mask from the previous section. It is only required to create an array with identical dimension and then copying the complementing colour values. The complementing colour values are simply mirrored along the 127 axis (meaning a value of 140 will become 114). In code, both of the overlays are generated at the same time to be more efficient (not having to iterate multiple times).

Mask application

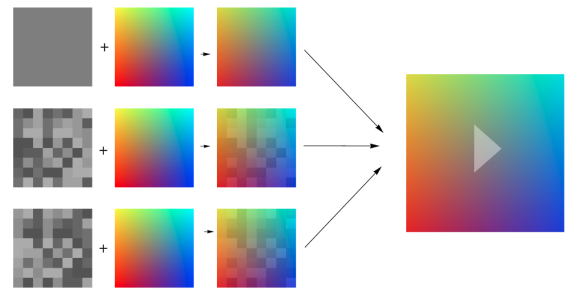
After generating all these masks, it is only required to overlay them over the original frames. First, transparency of the framemasks is set. The value of transparency is another aspect that needs to be selected. During development, it has however showed that it has a quite similar impact to shifting all the values closer towards the middle value - 127. This effect was further explored during the experiments (meaning it will be presented as one of the variables).

Afterwards, the semi-transparent framemasks are applied using alpha-composition. After this step, three frames exist - the original frame and the two data ones. At this point, if the two data frames were overlayed, the result should be the original frame.

Video generation

Using the previously generated frames, a video is created. This video runs at 120 frames per second. As there are only three frames in this video, this would mean the video lasts around 0.025 seconds. This is of course too short to use and therefore the video is looped to end up with a video of useful length. In practice, the desired length of the video depends greatly on the application however for testing, the video was looped to be around 20 seconds - details and reasons for this are explained in 4.4.

Figure 3: A simple illustration of the whole generation process



4 Experimental Setup and Results

In order for the results to be as unbiased and useful as possible, a rigorous testing procedure is needed. The following sections will outline the equipment used, the way the experiment was set up, the experimentation process as well as analyzing the results and discussing measures that had to be taken due to time constraints that could limit the usefulness of the data.

4.1 Experimental environment

All of the experiments were ran using the code on the Github branch - [4]. The experiments ran on two monitors, the first being Iiyama GB3461WQSU-B1 - a monitor with poor color accuracy but a high refresh rate (144Hz) and the second being Dell U2515H, a monitor with decent color accuracy but only a 60Hz screen. The camera used was the 48MP main camera of Asus Zenfone 6 - ran at 1080p@240fps.

4.2 Environmental elements

The previous section described the monitors and the camera used for the experiments. While this should allow the reader to reproduce the results, the conditions of the experimentation were rather specific and may not be representative of the circumstances within which this technology would be used in the real world. To combat this, the following sections will examine specific attributes of the used technologies and their impact. This will allow the reader to form a rough image of capabilities of this technology in common environments.

Display colour accuracy

As recognizing small nuances between different colours is the cornerstone of this technology, it is of paramount importance for the screens to be as colour accurate as possible. In reality, the colour accuracy of most displays is quite questionable. The colour accuracy of displays varies even between different monitors of the same model. While specific measures of the used monitors are not provided, the monitors can be taken as representative examples of two classes. The Iiyama monitor is one of the cheapest ultrawide monitors with a 144Hz refresh rate, and as such, the colour accuracy was not the main focus. Therefore, the monitor is a representative of the ”cheap” high refresh rate monitors. The second monitor used (Dell) is a representative of a midrange 60Hz monitors. This class of monitors is among the most commonly used by regular users. A decent colour accuracy is to be expected, as these monitors are usually well rounded (nothing major about them to warrant a big part of the cost).

Maximum display frequency

One of the most limiting requirements of this technology is the minimum refresh rate needed to establish a seamless encoding. As testing proved, there currently is not a way to encode data on a 60Hz monitor without visible artifacts. The situation is better on a 144Hz monitor. It is unclear whether a higher refresh rate would lead to even better results however considering the rarity of even 144Hz monitors, this will not be very relevant for daily use.

Camera detail

As the focus of this paper is on screen-to-camera communication, the camera aspect is logically one of the most important ones. For testing, a simple average phone camera was used - Asus ZenFone 6. This is a rough estimate of what can be used in real world. However considering the speed at which the phone cameras are evolving, these measurements may not be as relevant in the coming years. Should there be any future research (ideally with more cameras), it may be possible to create some sort of a timeline at which the technology improves with regards to screen-to-camera communication.

4.3 Variables generation

The previous chapter included a thorough examination of the aspects that affect the encoding as well as the effect they would have on the colour range. This section will outline how and what values were generated and how they were intersected to able to be used in testing.

Colour range width

As mentioned, colour range width is one of the most important aspects. A wider colour range allows to reliably transmit several colours (meaning a decent data throughput) however if it is too wide, the encoding will become visible. Therefore, several different values were created to test the maximum width possible - hoping an ideal value can be found which would allow for as much transmission of data as possible with no observable artifacts.

1.	2.	3.	4.	5.	6.
2	6	12	15	20	40

Figure 4: List of values used to test colour range width

Colour width

Colour width or number of colours denotes how many different colours can be located in a colour range. A smaller width means more colours which leads to higher data rate however that makes it harder to distinguish between colours. This will be tested by the camera. Following values were used in the experiment.

1.	2.	3.	4.	5.	6.
1	2	4	5	6	10

Figure 5: List of values used to test colour width

Minimum distance from the original

This denotes the distance the colour closest to the 127 axis will be. While this distance effectively subtracts from maximum possible width of the colour range, it is necessary to make sure that the colours differ enough from the original image for reliable decoding.

1.	2.	3.	4.	5.
1	3	5	10	15

Figure 6: List of values used to test minimum distance from the original.

Transparency

The overlay masks detailed in 3.3 are applied using alpha composition [6]. For this, they need to have a certain level of transparency. During development, it was found that generally higher transparency has a similar effect to reducing the overall length.

1.	2.	3.
32	48	64

Figure 7: List of values used to test minimum distance from the original.

Variable combination

Combining different values from previous sections, a table was created to test different configurations. These configurations start small - with low likelihood of being observed and gradually increase to values that should hopefully guarantee data transmission but at a significant risk of being observed.

4.4 Experiment procedure

Due to time related circumstances, the extent of the experiment had to be reduced severely. The exact reasons are explained in sectrion 5 however in short, there was not enough time for the due process to ensure the best interests of possible research participants. An approach that would make the data more objective is proposed in section 7.2. This section will explain the process that was done instead.

Even though this technology operates under the assumption of some ”standard” color-accuracy of the human eye, the reality is that every person sees a bit differently. To measure the effect of this, the author undertook a graphical aptitude test before doing any experimentation (described in detail in

Configuration #	C.R.W.	C.W.	M.D.	Transparency
1.	2	1	1	32
2.	2	1	10	32
3.	6	2	3	32
4.	6	2	10	32
5.	12	4	5	48
6.	12	6	10	48
7.	15	5	15	48
8.	20	10	5	64
9.	20	10	10	64
10.	40	10	15	64

Figure 8: A table of all used configurations. C.R.W. refers to colour range width, C.W. to colour width and M.D. is minimum distance from the original

4.5). The second step was test preparation. While the encoded pictures were generated using the values explained in 4.3, it was also crucial to make sure the person undertaking the test - the author - does not know the right answer to prevent bias. Therefore, the generated files were randomly prepared by a third party. The participant noted down the answers while completing the test and later obtained the right answer from the third party.

Detailed experimentation walkthrough

As mentioned in the previous sections, the experiment started with generation and randomization of encoded images. At the end of this step, ten folders (one for each configuration) were created, each containing a set of five pairs of encoded and original images. The author looked at each pair, try to answer the question: "Which one is the original image?". After noting down the answer (within the time limit of 20 seconds), next pair of images was presented. After analyzing all 50 pairs, the test was rerun the same way (but with the folders rerandomized) on the other monitor. After the second round of testing, the user answers were compared to the correct answers.

4.5 Graphical aptitude test

As the technology is based on using slightly different shades of colours hiding information in images, it is most useful to test the user ability to distinguish between closely related colours.

There are several tests online that are designed to do just this. They are most often called color perception tests and for the purposes of this paper can be divided into two categories - first being self reporting tests where the user sees a spectrum and is supposed to report the number of colours they see. This is quite prone to issues as users "know" the right answer (the answer width is standardised and most users will figure out the number of possible columns). They can therefore consciously or unconsciously cheat. The other kind is more useful for the purposes of the paper as the measurements are usually more objective. The user is shown a list of colours arranged by gradient with one hue missing. The user then has to pick the right hue from a list of options. There is no way to cheat on this kind of test apart from random guessing which will be considered when analyzing the results.

The final test used by this paper will be based on the approach described in the second paragraph. It is a Farnsworth Munsell 100 Hue Test [7] based test provided on the site [8]. This test provides a score on a scale 0-100+. Score below 20 can be considered very good, with average being between that and 100. Anything above 100 is quite poor. While this test is an industry standard, for it to be fully effective it needs to be completed in a controlled environment. Again, due to time constraints, this was not possible. It is also noteworthy, that if the tests were to be conducted with a number of participants, it would be needed to put much more thought into this part - using a third party test could lead to unwanted sharing of privileged information.

4.6 Experiment log

The testing did not go as planned. Starting with the graphical aptitude test, which the author received perfect score in. This is likely due to luck (guessing the right colours) and the experience acquired while researching these methods and the topic itself. Therefore, the results obtained can be considered a worst-case scenario for the human visibility (the author notices even small artifacts created by encoding) and best-case scenario for transmission (the small differences between colours that would allow to transmit several bits are observed).

After this start, the experiment ran once. After this run was finished and analysed, a bug in the software was identified which reduced the number of colours within the encoding. Therefore, the entire experiment had to be reran. After conducting the now fixed experiment, the selected method for evaluation proved to be quite ineffective - with certain encodings, there was no discernible difference between images in the pair and the results were therefore pure guesses. With others, the difference was immediately obvious and the author was able to guess the encoded image with 100% accuracy. Therefore an alternative grading system based on the observed differences between configurations was created. While virtually all (all that could transmit any data) encoded images were correctly identified, certain artifacts were more visible than others. Details of this system are explained in the following section. Two more categories are considered - one which considers whether the grid is recognizable on a picture taken by the camera and the other how many of the possible colours were recognized.

Observability

This score denotes on a scale of zero to ten how noticeable the flickering/visual artifacts are. Zero denotes absolute invisibility and ten very noticeable (noticeable by everyone even if far from a monitor). The scores between these values were assigned based on how much focus was needed to observe the flickering - lower scores require higher focus. This was done by unfocusing and squinting eyes till the flickering became invisible. Values on this scale should be relative to each other, meaning two configurations with same observability score should require the same amount of focus to observe. This score has issues with replicability and biases and is one of the weak points of the experiment. However, as mentioned before, with each eye a bit different, it is difficult to measure this objectively.

Grid recognition

This is a discrete score that is assigned to each configuration based on how reliably the grid can be observed. The following values can be assigned.

Score	Meaning
0	The grid is not visible
0	Parts of grid are visible in the center
2	The grid is visible in the center of the picture
3	The grid is mostly visible throughout the picture
4	The grid is reliably visible throughout the picture

Figure 9: A table outlining the scoring used to evaluate the results

Colour number

This is a score that denotes how many colours were visible (out of the all colours encoded) in the area denoted by the grid recognition score.

4.7 Results

Configuration #	Observability	G.R.	Colour #
1.	0	0	0(4)
2.	2	2	2(4)
3.	1	1	2(6)
4.	3	3	2(6)
5.	3	1	2(6)
6.	4	3	2(6)
7.	6	4	4(6)
8.	5	4	2-4(4)
9.	6	4	4(4)
10.	8	4	8(8)

Figure 10: Table containing the results of tests on the first screen - 144Hz monitor. G.R. stands for grid recognition

Configuration #	Observability	G.R.	Colour #
1.	0	0	0(4)
2.	3	2	2(4)
3.	1	1	2(6)
4.	4	3	2(6)
5.	4	1	2(6)
6.	5	3	2(6)
7.	7	4	4(6)
8.	6	4	2-4(4)
9.	6	4	4(4)
10.	9	4	8(8)

Figure 11: Table containing the results of tests on the second screen - 144Hz monitor. G.R. stands for grid recognition

4.8 General remarks

From the results of this experiment, several improvements are quite apparent. Firstly, in some cases, the biggest problem was low minimum distance/high transparency. In these cases, there was a significant drawback to the data transmission (sometimes being completely impossible) while still causing visual artifacts. It seems that reliable transmission of multiple colours is only possible with sufficient colour range and decent combination of transparency and minimum distance. This however leads to visual artifacts (which are even more noticeable on the lower frequency screen). The use cases of the technology is elaborated upon in section 6.

5 Responsible Research

As with any scientific endeavour, it is crucial to consider any possible ramifications the research might have. In the context of this project, this means reproducibility of research, the integrity and the ethical concerns of potential user research.

5.1 Reproducibility

While designing the experiment as well as when writing the papers, there was a significant emphasis on making the process as clear as possible and to provide all the variables needed to reproduce the results. Not only is this ethical and a good scientific practice, this also means that future contributions to this project will have a basis to build upon. This is especially relevant considering that a part of this project was human research - one that could not be conducted due to reasons explained in 5.3

5.2 Scientific integrity

As many other research papers, this one proposes an idea, tests it and later evaluates the results. While this paper showcases quite a few methods and approaches to prevent conscious or unconscious bias, there is always a way to modify the results to better fit the expected and desired outcome. As mentioned, unconscious bias was something that was largely accounted for - using randomization and blind tests. The conscious bias ultimately depends on the honesty of the researcher. However should there be any concerns about the validity of the results of this paper, the paper hopefully provides enough information for other researchers to repeat the testing and check the results - as further elaborated on in 5.1.

5.3 Ethical aspects of potential human research

While designing this research, there was originally a plan to conduct a user research that would help better determine the observability of the encoded data. It, however, proved unfeasible to conduct this research ethically within the allocated time period. The ethical guidelines required a full ethical checklist, as well as a detailed consent form and a data plan. The data plan proved most problematic, as there were potential data sharing issues with the chosen graphical aptitude test. Solving these issues would require re-implementing the test on a safe platform (possibly offline). Another theoretical issue was the health aspect of the research, as it may be possible to cause headaches or other similar symptoms via rapid flashing lights (experimentally this is a problem when flickering occurs at a frequency lower than 100 Hz but further research is required). Having considered all these issues, it was decided that a safer approach would be if only the author was conducting the experiment. This way, there can be no uncertainty about the informed consent about data handling or other possible adverse effects. For further conducting of this research, a set of loose guidelines based on the original plan is presented in the section 7.2.

6 Discussion

With the raw data being displayed in 4.7, it is important to put these results into a bit more context and consider potential uses for this technology. This section will explore all the trade-offs considered, possible issues of experimentation and finally use cases where the technology can be utilized even with current limitations.

6.1 Trade-offs

While the configurations don't cover all the possible combinations of values, the values chosen represent different approaches and try to measure at which point certain levels of transmission could be established (reliable single colour, reliable multi-colour etc). From the results it is however clear, that with higher possible data throughput, the observability also increases. This is an expected result, and it depends on the application that is desired. The technology is mostly useful in applications where high data throughput is desired and the observability is not such a concern. In places where it is crucial for this to remain hidden, unmodified approaches proposed by papers mentioned in the introduction are more useful.

6.2 Experimentation issues

While the experimentation procedure and the method of recording results was designed quite thoroughly, it relied on some false assumptions. The biggest problem was the clear distinguishability of encoded images. The chosen approach assumed the pictures would be hard to tell apart, but in all but one configuration, it was obvious which image was encoded. This also goes in hand with the second issue, that being the experience of the author. As the author never did any graphics related work in the past, it was assumed that the graphical aptitude test would yield an average score. However due to experience picked up during development and research of this paper, the author became quite experienced in noticing the small differences between different colours and was able to score a perfect score on the test. This resulted in the encoding being subject to higher scrutiny than they might be in the real world. Another perceived effect was encoding on the Dell 60Hz display. This had a smaller impact than expected, only increasing observability in some of the configurations. Therefore, it is quite clear that higher frame rate is much more important than color accuracy.

6.3 Possible applications of the technology

While the unfortunate conclusion is that this implementation of the technology is recognizable by a healthy, trained and focused eye, this does not necessarily mean that the technology cannot be commonly used. With a fine-tuned configuration (details specified in 7.2), it may be possible to achieve a reliable multi-colour transmission at a low observability level. Therefore, the technology could likely be used in information and advert panels in public areas - screens majority of people aren't looking at unless they are intrigued by them and would like to learn more. There is also a higher distance between the user and the screen - something that seems to reduce the observability of flickering during the testing (this was not rigorously tested). The proposed approach does not seem mature enough to be used with common household monitors - at least without significant visual artifacts.

7 Conclusions and Future Work

This sections will briefly conclude by answering the main research question of this paper as well as proposing further improvements and directions this research could head.

7.1 Conclusion

Answering the question: "Can multiple luminance levels increase the data throughput of screen-to-camera communication?" proved to be quite difficult. While it is definitely possible to reliably transmit data using several different colours (in the most ambitious configuration, this would lead to quadrupling the data throughput), this comes at a cost where the encoding is clearly visible. In the results sections, several viable configurations were examined. In the eyes of the author, the last three configurations are the most practical ones. These allow to reliably transmit increased amount of data. However considering the problems of the experiment (detailed in 4.6 and potential future improvements (7.2, it is quite likely that using multiple colours may be possible at a lower observability value.

7.2 Future work

As explained in previous sections, during this research, a lot of possibilities for improvements were observed. This sections will outline the major ones and explain other directions this research can head.

More objective research

Several corners had to be cut during experimentation due to ethical and time constraints. A new research, building from some conclusions learned in this paper (using them to test different configurations) could figure out new and better approaches to transmit the data as well as provide a more realistic observability score (or even design a more objective measure). As mentioned in this paper, this was originally the plan and some of the steps mentioned in 4 can be reused - graphical aptitude testing and picture randomization steps. There are however reasons why this method was not originally - requiring additional time and effort to handle all ethical aspects, associated with such testing (for reasons explained in 5).

Configuration fine-tuning

As the insight into the effects of the variables was mostly theoretical, the configurations were chosen to explore vastly different options to determine few decent ones. Consequently, it should be possible to fine-tune the configurations to a point where the observability is as low as can be while being able to reliably transmit multiple colours. The author approximates, that it should be possible to achieve four colour transmission at observability levels of four or maybe even three.

Smart encoding protocol

During testing it was observed, that it may be beneficent to cluster squares of similar colours together. This makes it easier to distinguish the individual squares as well as accurately tell their colours (as they can be compared to a similar square right next to it). Therefore, a smarter encoding algorithm may be able to encode grids in a way where it is possibly to more reliably use lower observability configurations. It may also be noteworthy to explore differently sized grids (the testing ran on a grid where each square was 32x32 pixels) and possibly even grids with alternate patterns (not only squares).

References

[1] Alan A., Bill B., and Cathy. Title of article. *Title of Journal*, 10(2):49–53, 2005.

[2] Tianxing Li, Chuankai An, Xinran Xiao, Andrew T. Campbell, and Xia Zhou. Real-time screen-camera communication behind any scene. May 2015.

[3] David L. MacAdam. Visual sensitivities to color differences in daylight. 1942.

[4] V.Crha. Luminance level variation. https://github.com/VojtechCrha/luminance-level-variation, 2022.

[5] A. Wang, Z. Li, C. Peng, G. Shen, G. Fang, and B. Zeng. Inframe++: Achieve simultaneous screen-human viewing and hidden screen-camera communication. pages 181–195, 2015.

[6] Wikipedia. Alpha compositing. 2022.

[7] Wikipedia. Farnsworth–munsell 100 hue test, 2022.

[8] X-Rite. Graphical aptitude test, 2022.

[9] Kai Zhang, Chenshu Wu, Chaofan Yang, Yi Zhao, Kehong Huang, Chunyi Peng, Yunhao Liu, and Zheng Yang. ChromaCode. October 2018.

[10] Z Zhao, P Grosso, R Koning, J van der Ham, C de Laat, M Ganzha, and M Paprzycki. 2010.