tagAR

Auditory and haptic feedback to make interaction more natural and satisfying

Group 1
Erik Lindberg
Moa Winberg
Patric Hjorth
Sebastian Blomkvist

DT2140 Multimodal interaction and interfaces

Abstract
Multimodal digital interfaces are useful when emulating activities from the real world. This project sets out to explore if interaction with an augmented reality application can be made more natural and satisfying by providing the user with auditory and haptic feedback. To examine this, three versions of an augmented reality graffiti application was developed; one with (aside from visual feedback) auditory feedback, one with haptic feedback, and one with both haptic and auditory feedback. How easy, natural, and satisfactory users perceived the interaction with the different versions of the application was then evaluated. The results of the evaluation indicated that the version with both haptic and auditory feedback, complementing visual feedback, made interaction more natural, satisfactory, and easy; i.e. the more redundant feedback modalities which occurs in the natural activity is mimicked, the better. Furthermore, auditory feedback made a greater positive impact in all categories than haptic feedback in this context, which may be due to sound being more salient than haptics in painting graffiti.

Video Demo
https://vimeo.com/386720362

Introduction
Graffiti is an artform favoured by artist and counter-culture movements because of the free and expressive aspects of the medium. To a graffiti artist, the whole world can more or less be considered a canvas on which to display your message. Augmented Reality (AR) is a useful tool when emulating a real world activity, as it provides the ability to superimpose digital artefacts on top of a live video (van Krevelen & Poelman, 2010).

We wanted to emulate and elevate the experience of graffiti even more with the help of digital tools. By removing the element of risk we want to make it available to a wider audience, as well as having a future opportunity of bringing features not possible to this artform without the technological element. Therefore we decided to create a mobile AR application, emulating a satisfying digital graffiti experience.

Related work
A number of applications have been developed in the area of graffiti and augmented reality. For instance, *Digital Sprayers* is an application which “turns your phone into a spray can” and allows users to express themselves through creating tags and graffiti art in augmented reality. *MarkAR,* another app, also turns your phone into a spray can, but have a social media component as well. Besides enabling users to create virtual graffiti in AR, the app will also let them share their artworks and stencils directly in the application with their friends and family. Another option of digital graffiti is *AR Graffiti Artist*. However, the usability of the application is highly questionable. You have to be outside to use it, but doesn’t even work well then. Additionally, the “spray cans” costs real money to use. Thus, not a good AR graffiti option. Google has developed an experimental AR art application called *Just a line*, which lets users create 3D artworks in open spaces, not only on surfaces. Instead of using

movements with the smartphone to paint, like an actual spray can, touch inputs on the screen is used to “spray”.

Using different feedback modalities and gestures as to make interaction with a device more natural has been explored in a number of studies. Auditory feedback is a natural way to notify users of their gestures and actions, especially for portable devices such as smartphones (Rocchesso & Bresin, 2007). However, the sounds should be tightly embedded into physical objects and controlled by a continuous flow of interaction in order to move away from schizophonia, i.e. the separation between the medium and the source, which diminishes the soundscape (Rocchesso & Bresin, 2007). Haptic feedback has also been utilized in order to increase user satisfaction. Hwang et al. (2017) enhanced the interaction with a virtual-reality piano, called AirPiano, by implementing mid-air haptic feedback. When a piano key was pressed, the feedback was used to mimic the natural resisting force of a real piano key. Their evaluation showed that the haptic feedback greatly increased user enjoyment and satisfaction, when compared to not using it.

Position and orientation of the device for interaction, such as moving a smartphone around to aim where to spray paint in augmented reality, may be a bit less accurate and may take a bit longer to input, e.g. compared to touch interaction, but is more intuitive and enhances the experience (Hürst et al. 2011). Using natural gestures as an input modality in augmented reality applications is a highly intuitive interaction method (Piumsomboon et al. 2013). Additionally, studies has shown that adding a complementary feedback modality to visual feedback in an interactive system boosts user performance (Burke, 2006; Prewett 2006). Although, different modalities have different strengths and weaknesses. The combination of visual-audio feedback is especially effective in a single-task context (Burke, 2006), while visual-haptic feedback is more suitable when multiple tasks are performed (Prewett 2006).

Research question

To what degree can interaction with an augmented reality smartphone application be made more natural and satisfying with auditory and haptic feedback?

Hypothesis

When modeling a natural occurrence in a digital application, the more redundant modalities that can be mimicked to this natural setting, the more satisfied and natural experience the user will have.

Limitations

We will not develop any back end for the application or store anything locally, i.e. no information will be saved between sessions. This because it would take too much time to implement and it is not one of the core tasks needed to perform the user tests. We will not try to combine graffiti or street art with digital media (3D models, animations, videos, sound etc), but only push to simulate the analogue experience of conventional means of graffiti.

Methods

App development

The application was developed with the language Swift (version 5.1) in Apple's program Xcode. The AR environment was enabled using the software development kit ARKit 3.

The AR scene was added to the application using ARKit. The digital spray in the application consisted of an image placed on a 2D plane object. When the user pressed the spray-button, the object was attached to the scene on the nearest point (feature point) on a plane surface found in the view. The feature points were found by using the built-in method called *hitTest(_:with:).* In order to get a more continuous flow of the spray, the function was executed repeatedly. The sound of the spray consisted of a mp3 file. The file was played using a package called AudioToolbox as long as the user presses the spray button. In the version where haptics is added, vibration from the package AVFoundation was played once a spray object was added to a feature point.

Evaluation

The evaluation set out to evaluate the satisfactorily aspect of the application, i.e. was the user satisfied when using it, and whether or not the application was easily used (Rubin, 2008). If more modalities makes the experience more natural, satisfaction of using the application should be higher. Applications which provides satisfaction usually enhances user performance (Rubin, 2008). The evaluation was carried out as a within-subject design, meaning that we had to take the transferred learning effect between the different tests into account (Rubin, 2008). To mitigate this side-effect we used counterbalancing, where we balanced out the order of which the participants tested the different versions of the application.

The test was then carried out in a well lit room and had 7 participants which were all students at KTH and between the age of 22 and 28. The participants performed each task on a wall. Before the test we surveyed them regarding their age, experience with AR, as well as graffiti and its surrounding culture. Each participant completed three tasks, each task with a different version of the application containing different compositions of modalities; Visual + Haptics + Auditory (VHA), Visual + Auditory (VA), as well as Visual + Haptics (VH). The auditory feedback was played directly from the smartphone's speakers. After each task, the participant answered three questions regarding the satisfaction and ease of using each version of the application. All answers were in a scale from 1 (strongly disagree) to 7 (strongly agree) In the first task the participant was instructed to paint a smiley face, in the second to write the letter "R", and the third to write the word "BAM".

Results

The application

The UI consisted of two buttons and a view of the camera, the 'AR scene' (see Figure 1). The first button was round and positioned at the bottom of the app. Long pressing this button would result in a spray in the center of the screen, following the gestures of the device.

Figure 1 - The tagAR-application. “TAGAR” written within the application.

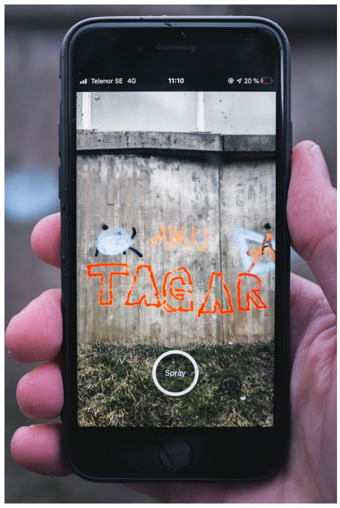

Figure 2 - Correct orientation

In order to be able to save creations without storing the data in a database, an option of taking a screenshot was implemented. By pressing a button on the lower right side of the screen (visible above the round button in Figure 1), a photo was taken and saved to the user's camera roll in the device.

The evaluation

The evaluation showed that the VHA version of the application was perceived by participants as being more satisfying to use, felt more natural, and that it made the task easier to complete (see Table 1). The VA version was rated marginally lower than the VHA version on all questions. On the other hand, the VH version was rated substantially lower in feeling satisfying and completing task with ease, but the same as the VA version in feeling natural. However, the results of this evaluation are not statistically significant and should only be viewed as indicative.

Version	**Complete task with ease**	**Felt satisfying**	**Felt natural**
VHA	5	5,4	5
VA	4,9	5,1	4,6
VH	4,4	4,4	4,6

Table 1 - Average answers from all participants when using the different versions

Discussion

App development

Different alternatives to platforms and languages were discussed before starting the development. Our supervisor suggested us to use the development platform Unity together with the language C#. Swift was eventually chosen because of its simple syntax and our desire to learn more about the language for future projects. None of us had implemented in Swift before but we had tried AR in Unity and C# in a former lab in the course. The application did not require any specific graphical assets, only one or two buttons, this also lead us to use Swift and Xcode instead of Unity. Another criteria was that we needed to be able to test the application in an easy way on our own iOS devices. Xcode allowed this without the need for additional programs.

The round button used for generating the spray was chosen in order to make the interaction natural to the user. In other apps such as Instagram and Snapchat, this circle indicates recording when long pressing, same interaction as we would use. The position and color of the button should also be familiar to the user and the button should be easy to reach with your thumb.

Evaluation results

The results from the evaluation is in many ways in line with previous studies. Having additional feedback modalities, such as auditory (Rocchesso & Bresin, 2007), complementing visual feedback should make interaction feel more natural, satisfying, and easier for the user (Burke, 2006; Hwang et al., 2017), and our evaluation supports this claim. Additionally, the more feedback modalities, which are present in the natural occurrence of the action, that are added, the better. However, adding only auditory feedback made a more positive impact than adding only haptic feedback to the application. This could perceivably be because auditory feedback makes a greater positive impact on performance while performing a single task, when compared to haptic feedback, and this evaluation solely contained single-task tasks for the participants to perform. Why haptic feedback did not have a greater positive impact, such as suggested by Hwang et al. (2017), could be because haptic feedback is not as salient when using a spray can, which is what the application of this project tries to mimic, compared to playing the piano. Furthermore, our evaluation did not evaluate a version of the app not using multimodal feedback, so it is not determined how the VH version would fare compared to a uni-modal version only using visual feedback.

Evaluation design

KTH students would hopefully be a selection of participants with a higher than average familiarity with AR-applications. This familiarity should be beneficial for this study in two ways. Firstly, they may be a potential user group for the AR-application. Additionally, if the participants would be new to AR the new experience might overwhelm the participants and changes in feedback modalities may not be perceived and go unnoticed. Being familiar with AR-applications would counteract this. However, even though the participant in some cases did not perceive, at least consciously, any difference in the feedback modalities between the versions of the applications, the version with auditory and haptic feedback was still rated the highest in satisfaction, ease of use, and feeling natural.

As to not make the application, and hence the smartphone, to be perceived as a schizophonic artifact it was decided to refrain from the option to have the participants of the evaluation receive the auditory feedback through headphones. Thus, uncoupling the audio from the source. Instead, the auditory feedback was directly conveyed via the smartphone's speakers, which would hopefully have a more positive impact on the soundscape. Since the versions of the application which utilized auditory feedback achieved a better result in the evaluation, i.e. it was more satisfying, natural and easy to use, than the one that did not utilize it, it may be concluded that the hoped-for impact was produced.

Method critique and future work

The application could be further developed so that the feature points that the spray attach to were more accurate to the objects in the scene. Sometimes it occured that the points changed position and did not stick to a plane surface because a feature point with nearer or farther distance were found. The spray could also only be seen from one direction since it was in 2D and could not recalibrate by its own. This meant that a user could not change direction of the device while spraying. Since the test were only to be perform at one specific surface, this was not necessary at this point but in a different set up, it would be. This could be done for instance by adding a method for calibrating the device that would feel natural to the user, or by using a different method than attaching 2D plane objects. A different method could be to create a 3D line instead and then normalizing the points to the surface.

As a result of the aforementioned issues of accuracy, some problems arose during the evaluation when participants performed the tasks. We had not prepared how and to what extent we would aid the participant when technical problems obstructed their ability to perform the task, and we felt incentivised to intervene in some manner, so that the problems would not interfere with how they would perceive the feedback, thus impacting our results. In order to be better prepared we could have conducted pilot studies, allowing us to earlier identify these issues and better plan for how we should intervene.

Furthermore, more interactions with the user could be added, for instance the ability to change color of the spray can. It could also be of interest to evaluate how a different input modality would impact satisfaction. Instead of using the point-and-shoot like feature to select where to spray, an alternative version could use touch input to select where to spray, i.e. you would paint with your finger. This would percievably be an easier input method, but less intuitive (Hürst et al. 2011), and an evaluation could determine whether it would lead to increased user satisfaction or not, however possibly at the cost of feeling less natural. To further explore if a greater amount of feedback modalities leads to better user experience, a future evaluation could include feedback modalities which does not occur naturally in the activity.

An additional flaw with the current evaluation, as previously mentioned, is that it does not evaluate a uni-modal version of the application. This would provide us with a better baseline from which the impact of the different modalities could be measured. For instance, rather than only comparing the VA and VH versions to each other, their impact could be measured on the basis of a uni-modal version. Instead, we currently only look at how different combinations of feedback modalities affect interaction and not how multimodal feedback affects interaction when compared to uni-modal feedback. Hence, we can not conclude that multimodal feedback provides a more natural and satisfying user experience, than uni-modal feedback. Due to the limited nature of the study, the results are only indicative. To be statistically significant more participants would be needed in the study and the app would need to be more consistent.

Conclusion

This study has yielded indicative results which suggests that users will perceive a difference in how satisfactory, natural, and easy interaction with an AR-application is depending on which combination of feedback modalities the application utilizes. To summarize; the combination of visual, haptics, and auditory feedback made interaction the most satisfactory, natural, and easy; visual and auditory feedback was perceived marginally worse in all categories; and visual and haptic feedback was perceived as the worst. Auditory feedback may have been a better addition to the application than haptic feedback due to it being a more salient modality in the natural occurrence of the activity or because it enhanced task performance to a greater extent. In further studies, more interactions with the user could be added as well as feedback from modalities that do not occur naturally in the activity. This could explore if increased number of modalities improves the user experience or not.

References

Burke, J. L., Prewett, M. S., Gray, A. A., Yang, L., Stilson, F. R. B., Coovert, M. D., Elliot, L. R., Redden, E., (2006). Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. In: Proceedings of the 8th international conference on Multimodal interfaces (ICMI '06). ACM, New York, NY, USA, 108-117.

Gwilt I. (2018) Augmented Reality Graffiti and Street Art. In: Geroimenko V. (eds) Augmented Reality Art. Springer Series on Cultural Computing. Springer, Cham

Hürst W., van Wezel C. (2011) Multimodal Interaction Concepts for Mobile Augmented Reality Applications. In: Lee KT., Tsai WH., Liao HY.M., Chen T., Hsieh JW., Tseng CC. (eds) Advances in Multimedia Modeling. MMM 2011. Lecture Notes in Computer Science, vol 6524. Springer, Berlin, Heidelberg

Hwang I., Son H. and Kim J. R. (2017) Air Piano: Enhancing music playing experience in virtual reality with mid-air haptic feedback. In: 2017 IEEE World Haptics Conference (WHC), Munich, pp. 213-218.

van Krevelen, D.W.F., & Poelman, R. (2010). A Survey of Augmented Reality Technologies, Applications and Limitations. The International Journal of Virtual Reality, 9, 1–20.

Piumsomboon T., Clark A., Billinghurst M., Cockburn A. (2013) User-Defined Gestures for Augmented Reality. In: Kotzé P., Marsden G., Lindgaard G., Wesson J., Winckler M. (eds) Human-Computer Interaction – INTERACT 2013. INTERACT 2013. Lecture Notes in Computer Science, vol 8118. Springer, Berlin, Heidelberg

Prewett, M. S., Burke, J. L., Gray, A. A., Yang, L., Stilson, F. R. B., Coovert, M. D., Elliot, L. R., Redden, E., (2006). The benefits of multimodal information: a meta-analysis comparing visual and visual-tactile feedback. In: Proceedings of the 8th international conference on Multimodal interfaces (ICMI '06). ACM, New York, NY, USA, 333-338.

Rocchesso, D., & Bresin, R. (2007). Emerging sounds for disappearing computers. In Streitz, N., Kameas, A., & Mavrommati, I. (Eds.), The Disappearing Computer (pp. 233-254). Berlin Heidelberg: Springer.

Rubin, J., Chisnell, D., and Spool, J., (2008). Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests, second edition. New York, NY, USA: John Wiley & Sons, Inc.