Automatic Adjustment of Stereoscopic Content for Long-Range Projections in Outdoor Areas

Behnam Maneshgar
Immersive and Creative Technologies Lab,
Dept. of Computer Science and Software Engineering
Concordia University
Montreal, Quebec, Canada

Leila Sujir
Elastic Spaces Lab,
Dept. of Studio Arts
Concordia University
Montreal, Quebec, Canada

Sudhir Mudur
3D Graphics and Virtual Worlds Lab,
Dept. of Computer Science and Software Engineering
Concordia University
Montreal, Quebec, Canada

Charalambos Poullis
Immersive and Creative Technologies Lab,
Dept. of Computer Science and Software Engineering
Concordia University
Montreal, Quebec, Canada

ABSTRACT

Projecting stereoscopic content onto large general outdoor surfaces, say building facades, presents many challenges to be overcome, particularly when using red-cyan anaglyph stereo representation, so that as accurate as possible colour and depth perception can still be achieved.

In this paper, we address the challenges relating to long-range projection mapping of stereoscopic content in outdoor areas and present a complete framework for the automatic adjustment of the content to compensate for any adverse projection surface behaviour. We formulate the problem of modeling the projection surface into one of simultaneous recovery of shape and appearance. Our system is composed of two standard fixed cameras, a long range fixed projector, and a roving video camera for multi-view capture. The overall computational framework comprises of four modules: calibration of a long-range vision system using the structure from motion technique, dense 3D reconstruction of projection surface from calibrated camera images, modeling the light behaviour of the projection surface using roving camera images and, iterative adjustment of the stereoscopic content. In addition to cleverly adapting some of the established computer vision techniques, the system design we present is distinct from previous work. The proposed framework has been tested in real-world applications with two non-trivial user experience studies and the results reported show considerable improvements in the quality of 3D depth and colour perceived by human participants.

MM'17, , October 23–27, 2017, Mountain View, CA, USA.

ACM ISBN 978-1-4503-4906-2/17/10…$15.00
https://doi.org/10.1145/3123266.3123269

CCS CONCEPTS

• **Human-centered computing** → **Visualization systems and tools**; **Visualization design and evaluation methods**; • **Computing methodologies** → *Camera calibration*; *3D imaging*; *Shape inference*; *Reconstruction*;

KEYWORDS

stereoscopic compensation;stereoscopic projection;outdoor projection mapping;long-range projection;radiometric compensation;smart projection

1 INTRODUCTION

Many methods have already been proposed for generating stereoscopic content, perhaps the most popular of which is the anaglyph 3D. In contrast with passive and active vision technologies, anaglyph 3D does not impose any additional requirements such as the number of projectors needed [as in the case of polarized stereo or active stereo which require two], expensive active vision/shutter glasses which also have a limited usage time-span due to recharging requirement, or expensive passive stereo glasses with polarized optics.

Hence, in comparison with the other available techniques, it is no surprise that the anaglyph 3D is still the most popular technique for creating stereoscopic content, primarily due to its simplicity and cost-effectiveness. The creation of an anaglyph 3D involves the encoding of two images, one for each eye, with different colors. Historically, the colors which have been used are red and cyan. The viewer can then perceive 3D with the use of anaglyph glasses: glasses which have one red and one cyan lens. Each colored-lens allows only the image which is encoded with the same color to pass through therefore ensuring that only one image is seen by each eye. The brain, and in particular the visual cortex, fuses the two images seen by the eyes into perception of a three-dimensional scene.

Anaglyph stereoscopic content has been successfully used in demonstrating scientific results [15], visualizing maps [8] health [18], and also used as an easy way of presenting 3D content online [21]. More recently, researchers comprising artists, designers and computer scientists among others, have begun exploring the potential of stereoscopic technologies with artistic practices, and in particular

for outdoor projection mapping. This poses problems depending on the type of the surface used for projection due to the red/cyan color contained in the anaglyph stereoscopic content. For example, projecting an anaglyph stereoscopic content on a red wall will cause the cyan encoded image to appear as blackish, and therefore the viewer will not be able to perceive it as a three-dimensional scene.

Moreover projecting onto outdoor surfaces such as building facades, itself poses several challenges. Firstly, the projection surface may contain parts with complex reflectance properties or low reflectivity which may interfere with the projected content. Secondly, to be able to account for the occurrence of these types of adverse object properties, one has to capture the geometry and light behaviour properties of the projection surface, which typically requires use of a calibrated camera-projector system. The calibration of the camera-projector system becomes a non-trivial task because of the long-range. Standard procedures such as those described in [20, 22] require that all optical systems are focused on the calibration board/object. However, in the case of long-range outdoor projection the camera-projector system is focused on the projection surface which is located at a long distance away e.g. $> 15m$. Capturing images of the calibration board from such a distance leads to poor coverage within the image which in turn results in improper calibrations. On the other hand, capturing images by placing the calibration board at a shorter distance leads to blurry images which again results in poor calibrations.

In this paper, we address many of these challenges relating to long-range projection mapping of stereoscopic content in outdoor areas. We present a complete framework for the adjustment of the content based on automatic derivation of the projection surface geometry and light behaviour properties. We formulate the problem of projection surface modeling into one of simultaneous recovery of shape and appearance. Our system is composed of 2 standard fixed video cameras, a long range fixed projector, and a roving video camera for multi-view capture. While most of the time, we have used two fixed cameras, one on each side of the projector, in cases when we have too many holes in 3D reconstruction caused by self occlusions, we have experimented with a third camera on top, and it has worked very well. The overall computational framework comprises of four modules: (a) calibration of a long-range vision system without the need for calibration boards, (b) dense 3D reconstruction of projection surface from the multiple calibrated camera images, (c) modeling of the reflectance properties of the projection surface from roving camera images and, (d) the adjustment of the stereoscopic content in an iterative manner. In addition to cleverly adapting established computer vision techniques, our principal contribution is the system design which is distinct from previous work. The proposed framework has been implemented and has been tested in real-world applications with two non-trivial user experience studies in public places with stereo projection onto building facades after dark and participation by random people available or passing through the venue. The results are reported and show considerable improvement in the quality of 3D shape and colour perceived.

The paper is organized as follows: Section 2 gives an overview of the state-of-the-art in the area and Section 3 presents a technical overview of the proposed framework. In Section 4 we discuss the calibration of the camera-projector system when dealing with long-range projection mapping. Next, Section 5 describes the reconstruction of the projection surface/scene. The light behaviour properties are modelled to their best approximation as described in Section 6. Lastly, Section 7 presents an adjustment scheme for transforming the source images to generate projections which are perceptually closest to the intended images given the constraints imposed by the projection surface geometry and light behaviour properties. Illustrative experimental results and the user study reports are provided in Section 8, and conclusions and future work in Section 9.

2 RELATED WORK

Many different methods have already been proposed for controlling and altering the appearance of projections onto various surfaces the majority of which seem to work quite well for controlled environments. Below we provide a brief review of some of the earlier work which is closely related to the work being reported in this paper.

Grossberg et al. [10] presented a method to control the appearance on small objects by using one camera and one projector. In their work, the focus is primarily on the spectral responses, spatially varying fall-offs, and non-linear responses in the projector-camera system which creates a strong dependency between the camera view and camera response. The computed radiometric model of the system is then used to compute the compensated image.

In a similar approach, Aliaga et al. [2] addressed appearance editing on the object using a projector. Multiple projections are used to improve the resolution and compensate the images by reformulating the problem as one of constrained optimization. An elliptical Gaussian is used to model projector pixels and their interaction between projectors. In [5] Bimber et al. introduce a view-dependent stereoscopic projection for compensating distortions caused by the scene's structure. They present an elaborate process which involves computing of inverse light transport to create compensated images and demonstrate the success in controlled environments. In recent work by Ahmed et al. [1] multiple projectors are used to reproduce the appearance of an object. Their system has better black levels and less contrast compression. Several works have addressed [6, 7, 19, 23] dynamic and moving scenes. One or more projectors have been used to continuously compensate the projected image while the objects on the scene are moving. A perceptually-based object appearance modification has been attempted by Law et al. [13]. They partitioned the projection surface into patches based on the target appearance colors to make it appear as similar as possible to the target.

All the aforementioned techniques have been shown to perform satisfactorily in cases involving projections of non-stereoscopic content from short-ranges i.e. $\leq 3m$. The primary reason for this is the calibration procedure involved which for longer distances in uncontrolled environments produces poor results; both in terms of geometric and radiometric calibration.

Perhaps the closest work to that proposed here is in Maneshgar et al. [14]. They present a method suitable for long-range projections. The calibration problem is resolved by using phase shifting patterns to calibrate the out-of-focus system. Once the system is calibrated, they address the ill-posed problem of recovering the reflectance

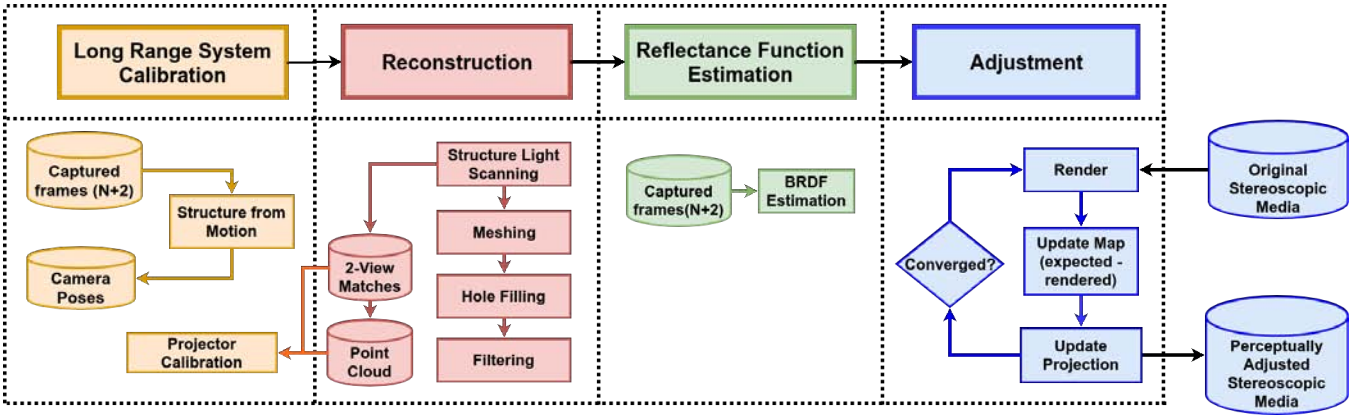

Figure 1: A diagram of the proposed framework summarizing the four processing modules.

properties from three samples and show how this is possible given the special characteristics of the application i.e. the audience is located close or near to the projector.

The work reported in this paper differs from all the above in that firstly the focus is long range and outdoor projection, and secondly, light behaviour properties of the projection surface are computed at pixel level resolution using the images from many different viewpoints provided by the roving camera, thus overcoming the sampling problem present in [14]. It cleverly adapts structure-from-motion (SfM) technique applied to multiple views from the different cameras to calibrate the two fixed cameras, thus avoiding all the problems associated with the use of a calibration board in long range projection. This makes the next step of recovering the geometry structure more robust. A novel method for compensation is also introduced by considering multiple bounces of each ray cast by the projector onto the 3D projection surface.

3 TECHNICAL OVERVIEW

The first step in the proposed framework is calibration of the system which involves the estimation of $N+2$ camera poses corresponding to the N frames captured by the roving camera and the 2 frames captured by the fixed cameras. Uniquely, our proposed approach leverages the sub-pixel accuracy of the dense reconstructions resulting from structured light scanning (SLS) techniques with the robustness and accuracy of the camera pose estimations resulting from Structure-from-Motion techniques (SfM). On the one hand, SLS techniques inherently involve a complex calibration procedure which becomes even more challenging as the number of cameras increases or the distance from the projection surface increases; this is due to the fact that the reconstruction can only be performed on the area visible to all the cameras i.e. the intersection area of all views. However, once the calibration is complete, an accurate and dense reconstruction can be generated. On the other hand, SfM techniques make only an assumption on the scene's rigidity and estimate the camera poses and intrinsic parameters which also can be used to create a sparse reconstruction of the scene, if necessary.

In this work, we overcome the difficulties of calibration and subsequent 3D reconstruction imposed by long range focus of multiple cameras. We do this by first computing the camera poses and intrinsic parameters using SfM on the set of $N+2$ images captured by the roving video camera and the two fixed cameras. In a second step, the two fixed cameras and a long range projector are used for SLS which yields us a dense reconstruction of the scene. This involves (i) projecting encoded Grey-coded patterns, (ii) capturing the patterns with the two cameras, (iii) decoding the patterns, and (iv) identifying the per-pixel matches. The per-pixel matches are then triangulated to produce a dense point cloud which is further processed to generate a dense mesh representing the projection surface.

With a fully calibrated system, the per-pixel reflectance function is estimated using non-linear optimization on the $N+2$ observations. This results in a reflectance map of the projection surface which is subsequently used for the adjustment of the images/videos to be projected.

Figure 1 shows a diagram of the proposed framework summarizing the four processing modules: long-range system calibration, reconstruction, reflectance function estimation, and adjustment. Each of these processing stages are described in greater detail in the following sections.

4 LONG-RANGE SYSTEM CALIBRATION

The long-range system calibration involves (a) estimating the camera poses and intrinsic parameters for $N+2$ views, and (b) calibrating the projector to the two fixed cameras.

4.1 Calibration of Long-range Cameras

Traditional calibration techniques such as the ones described in [20, 22] requiring multiple images of a known calibration object [usually a checkerboard] produce poor to no results when dealing with cameras focused at infinity, because of their strong assumption on the calibration board appearing in focus in the captured images. Therefore, using these techniques with a camera focused at infinity requires that the calibration board is positioned at a far distance [typically $> 15m$] from the camera, which leads to images in which the projection of the calibration board occupies a tiny fraction of the image's area. This in turn results in poor calibration parameter

estimations since the motion and intrinsic parameters [in particular distortion coefficients] cannot be accurately recovered.

To overcome these problems, we employ SfM for the calibration of the cameras. SfM involves capturing a large number of images from different viewpoints and performing feature extraction and matching with SIFT. As mentioned earlier, with the roving video camera swept randomly in front of the projection surface capturing the surface from different view points and view angles, and the two fixed cameras focused on the projection surface, we obtain $N + 2$ images. Bundle adjustment is then performed to simultaneously refine the parameters of the camera motion $[R|t]_{3x4}$, intrinsic camera parameters K_{3x3}, and the 3D points P in the scene while minimizing the re-projection error in all the images.

Figure 2: Long-range camera calibration. Traditional calibration techniques fail for long-range vision systems. In the proposed approach we use SfM for estimating the camera poses and intrinsic parameters, thereby eliminating the need for special calibration boards and procedures [3, 4]. Setup shown for experiment #1.

Figure 2 shows an example of the results obtained from SfM. Each camera is represented by its position and its oriented image plane. SfM generates a sparse reconstruction of the scene's structure which is also shown in the figure. However as we need a dense point cloud to accurately reconstruct the projection surface geometry, we do not use this sparse point cloud in further processing. Instead with these calibrated fixed cameras, we resort to the use of the SLS technique to obtain a dense 3D reconstruction.

4.2 Camera-Projector calibration

Once the cameras are calibrated, we proceed with the estimation of the projector's pose and intrinsic parameters, with respect to the calibrated cameras. Traditional techniques for camera-projector calibration such as [3, 4, 9, 16] suffer from similar problems, as previously mentioned, when applied to long-range systems. These techniques usually require projecting a pattern [usually a checkerboard] onto [or next to] the calibration board and detecting the corners in order to estimate the projector parameters. In other words, the projector is treated as an inverted-camera. However, projective geometry dictates that the farther away you move from the projector the larger the projected pattern. Hence, it becomes almost impossible to employ these techniques in long-range scenarios.

To overcome such problems, we employ SLS [11] for simultaneously (a) calibrating the projector with respect to the other cameras and (b) capturing the scene's geometry accurately. This process involves projecting encoded [with each pixel's location] patterns which are captured by the two fixed cameras and are then decoded to identify each pixel's location. Given the dense correspondences between the two images, a 3D position can be computed by triangulation. The result is a dense reconstruction of the scene's structure which in combination with the correspondences in the 2D projected image can be used to estimate the projector's pose and intrinsic parameters as in [22].

5 RECONSTRUCTION

As previously stated, SLS yields us a dense point cloud representation of the projection surface which is depicted as an XYZ map as shown in Figure 3a. Holes could result due to occlusions. These are filled with neighbourhood information. Bilateral filtering is used to remove noise if present in the mesh while at the same time preserving the edges. Finally, a mesh is created by triangulating the nearest-neighbours in the XYZ map, an example is shown in Figure 3b.

(a) **(b)**

Figure 3: Experiment #1: (a) The XYZ map of the scene's structure generated with SLS. (b) A render from a novel viewpoint of the reconstructed geometry.

The quality of the reconstructions using the above method were quantitatively evaluated by reconstructing known objects and comparing them with the ground truth. For objects located in a distance of $6m$ the estimated surface fitting error [RMSE] between the ground truth and the reconstructed object was in the order of $0.1cm^2$ which is comparable to the errors reported in [11].

6 LIGHT BEHAVIOUR ESTIMATION

We use the Bidirectional Reflectance Distribution Function (BRDF) to represent the light behaviour properties of the projection surface. The BRDF is estimated based on the per-pixel samples captured by the roving camera and the two-fixed cameras. Given the fact that during the projection mapping the audience is typically spread over a limited area in front or next to the projector's direction, we only consider the BRDF over a range of incident light and viewing directions ranging from $0 \leq \theta \leq \pi$ on the horizontal and $0 \leq \phi \leq \frac{\pi}{4}$ on the vertical.

We use the LaFortune BRDF [12] model in our work. This analytical model leverages the simplicity of the Phong model while

capturing realistic BRDFs from measured data. It should however be noted here that, although in this work we have employed this particular model, our method of fitting the measured data to a BRDF analytical model is not limited to it.

We use the $N+2$ per-pixel samples captured by the cameras to recover the LaFortune BRDF parameters via non-linear optimization. To sample the incident light from the limited area mentioned earlier, we perform a random walk using the roving video camera in front of the building (projection surface) so as to cover the surface and obtain adequate number of samples. During the estimation the per-pixel normals computed from the XYZ map are also incorporated. The result is the per-pixel BRDF; a diffuse map is shown in Figure 4a and the corresponding specular map is shown in Figure 4b. The inset picture in Figure 4b is a close-up of the region indicated with red and demonstrates the high level of detail we are able to capture using this approach.

(a) (b)

Figure 4: LaFortune BRDF: diffuse map(left), specular map(right). The inset picture in the specular map (b) is a close-up of the region indicated with red and demonstrates the level of detail captured by this approach.

7 AUTOMATIC ADJUSTMENT

The colors of the original image are adjusted such that any interference or colour distortions due to the surface's geometry and reflectance properties are mitigated to the extent possible. Using the recovered scene geometry, the surface's per-pixel BRDFs, and camera poses, an image is rendered from the projector's viewpoint and the colour difference is iteratively diminished. The steps are summarized in the algorithm snippet below.

$i \leftarrow 0$
$I^i_{rendered} \leftarrow Raytrace(K[R|t], I_{specular}, I_{diffuse}, I_{original})$
while $||I^i_{rendered} - I_{original}||^2 \geq \tau$ **do**
  $I^i_{update} \leftarrow I^i_{rendered} - I_original$
  $i \leftarrow i + 1$
end while
$I_{adjusted} \leftarrow I^i_{rendered}$

where $Raytrace(.)$ is a physics-based renderer [17] which given the [inverted] camera parameters $K[R|t]$ i.e. projector, the two maps containing the specular and diffuse coefficients for each surface point $I_{specular}, I_{diffuse}$, and the original images $I_{original}$ produces a render from a virtual camera [with identical pose and parameters as the projector] of how the original image will appear if projected onto the surface. During the computation the projector is modeled as an array of $N \times M$ point light sources each emitting light only to its corresponding surface point, where $N \times M$ is the size of the original image $I_{original}$.

The third column in Figure 5 shows an example of the adjusted image after a single iteration (top), and upon convergence of the algorithm (bottom). The iterative optimization is performed offline since for an image with resolution of 1600×1200 it typically takes 4 minutes to converge due to the rendering required at each iteration. Since this was more of a proof of concept, the computations have not been optimized, though clearly there is ample scope for it.

8 EVALUATION

The proposed framework was evaluated by human participants during two separate experiments. In both experiments we wanted to expose participants to the projection mapping of stereoscopic content and assess the effectiveness of the proposed approach with respect to color and depth perception. This was done through questionnaires which the participants had to answer before, during, and after the experiments. The research team conducting the experiments included four computer science researchers. The experiments received approval by the University Research Ethics Committee and informed consent forms were obtained from each of the participants. Furthermore, the participants were asked to fill out a demographic and background information form reporting on any prior experience or any dizziness problem with anaglyph 3D stereo, and were informed that they could withdraw from the experiment any time they felt so.

8.1 Experiment #1

The first experiment involved projection mapping of anaglyph 3D images on a shed. The shed was already installed as an exhibit in an indoor public place, the lobby of a building. The shed was chosen because of its red color which would interfere with the projection of the red/cyan stereoscopic content therefore causing problems with the color and depth perception of the viewers. The doors were white and provided a reference of minimum change.

In this experiment the projection was from a short-range [i.e. a little over 3 metres] because we specifically wanted to address only the following questions without introducing possible bias due to the distance of the projection:

- Does projection mapping of anaglyph 3D content on surfaces containing red or cyan colors cause interference with the color and depth perception of the viewer?
- Does the proposed approach improve the color and depth perception of the viewer?

8.1.1 Participants. This experiment involved 34 participants [70.6% male, 29.4% female] which according to the demographic form ranged from 19-39 years old [35.2% within the age group 25-26]. Of the 34 participants 44.4% had graduated [or were in the process of] from an art field, and 55.6% had graduated [or were in the process of] from an engineering field. These participants were random people who were around in the venue during the time the stereo projection was set up. The majority [55.8%] of the participants indicated that they had prior experience with anaglyph 3D and had no dizziness problem, although 2.9% of them could not perceive 3D.

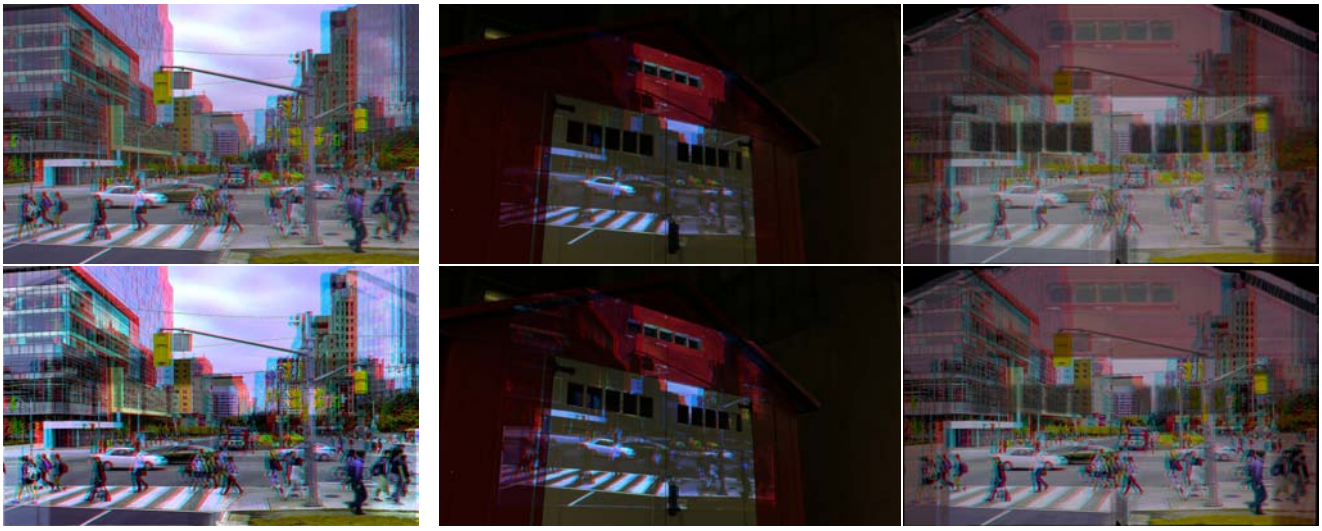

Figure 5: An example image used in Experiment #1. First column: (top) original image, (bottom) adjusted image. Second column: (top) projection of original image on scene, (bottom) projection of adjusted image on scene. Third column: (top) adjusted images after one iteration, (bottom) adjusted image upon convergence of the algorithm.

8.1.2 Setup and Equipment. As previously mentioned, this experiment involved projection mapping onto a red shed with white doors. The size of the projection surface was $1.35 \times 2.40m$ and the distance of the projector $3.3m$. A projector [SANYO PLC-ZM5000L] was used with a native resolution of $1920x1200$. The two-fixed cameras [Pointgrey Grasshopper 3] with a resolution of 1920×1200 were placed at a distance of 3.5. A third similar camera [Pointgrey Grasshopper 3] was used as the roving camera to capture over 200 images of the scene.

8.1.3 Procedure. The participants were provided with anaglyph 3D glasses so that they could observe the stereoscopic projection. Initially, the original non-compensated stereoscopic image was projected and after 10 seconds the adjusted stereoscopic image was projected. The participants were allowed to move freely in the general area of the projection during the experiment and observe the projection from different viewpoints.

8.1.4 Results. We present the qualitative and quantitative assessment of the results from this first experiment. Figure 5 shows the stereoscopic content before and after the application of the proposed approach. The first column shows the original non-compensated image (top) and the adjusted image (bottom). The middle column shows the projection of the original image (top), and the projection of the adjusted image (bottom) onto the scene. In the images in the first column the change is almost not noticeable, however as it is evident from the images in the second column there is a significant difference in the projections of the two. For example, the content projected onto the red areas of the scene is almost not distinguishable in the original image which results in very poor or no depth perception. On the other hand, the projection of the adjusted image appears brighter and perceptually correct in these areas. Finally, the third column shows the resulting image after a single iteration of the adjustment process (top), and the final adjusted image upon convergence of the algorithm.

Figure 6 shows a table of the statistical significance between the participants' questionnaire responses about any perceptual change in the projection before (vertical) and after (horizontal) the application of the proposed approach. The question for both projections was 'Rate your color and depth perception for this projected image/video'. A likert scale was used ranging from [1,10]. The table shows a high statistical significance for the improvement in the perception of the participants with the adjusted images. In particular, 50% of the participants who had rated the perception of the projection of the original image with a '3' had rated their perception of the adjusted image with a '5' which is higher than the average of '3'. Similarly, 45% with an initial rating of '6' had increased their rating to '8', 67% with an initial rating of '7' had increased their rating to '10', and 33% with an initial rating of '9' had increased their rating to '10'. In Figure 7 we show the distributions of the participants' responses for these questions.

8.2 Experiment #2

The second experiment took place in an outdoor public space at night and involved long-range projection onto a building's facade containing glass windows, complex decorative sculptures, carved columns, etc. with no control over the illumination conditions.

In this experiment the projection was from a distance of over 20 metres. Participants viewed the projection standing on the pavement of a downtown street. The experiment was conducted over two nights, as during the first night it was drizzling. A good number of people did watch the projection in spite of the rain. The experiment was designed to focus on the following:

- Is the proposed approach able to enhance the color and depth perception of the viewer for long-rage projections in outdoor areas?

Column %	5	6	7	8	9	10	NET
3	50% ▲	0%	0%	0%	8%	0%	6%
4	50%	0%	40%	18%	17%	0%	21%
5	0%	100%	20%	18%	8%	0%	15%
6	0%	0%	20%	45% ▲	17%	0%	24%
7	0%	0%	0%	9%	33%	67% ▲	21%
8	0%	0%	0%	9%	17%	0%	9%
9	0%	0%	20%	0%	0%	33% ▲	6%
NET	100%	100%	100%	100%	100%	100%	100%

Figure 6: Experiment #1: Statistical significance of participants' questionnaire responses between before (vertical) and after (horizontal) the application of the proposed approach. The question for both projections was 'Rate your color and depth perception for this projected image/video'. A likert scale was used ranging from [1,10]. 95% confidence level. Sample size: 34.

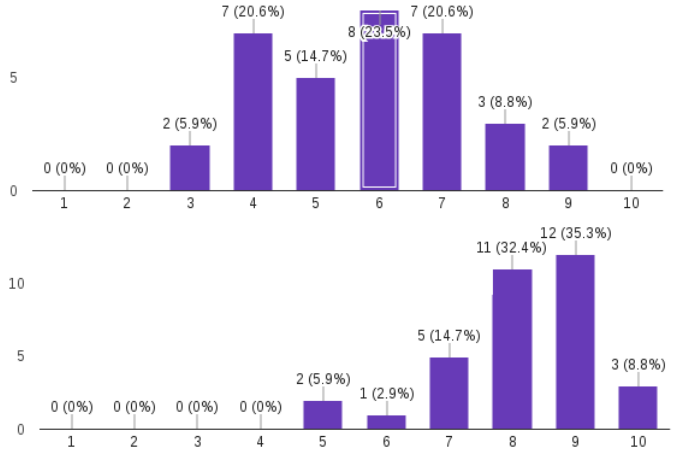

Figure 7: The distributions of the responses to the question 'Rate your color and depth perception for this projected image/video' with the original image (top) and adjusted image (bottom)

8.2.1 Participants. The second experiment involved 37 participants [64.9% male, 35.1%] over two days chosen at random from the street, which according to the demographic range from 18-68 years of age [62.1% within the age group of 22-29]. Of the 37 participants 73% had an engineering, 16.2% arts, 8.1% social sciences background. One participant reported he/she had suffered or was prone to epilepsy or seizure and did not take part in the experiment. 78.4% reported that they were able to perceive depth when using anaglyph 3D glasses in the past.

8.2.2 Setup and Equipment. The experiment involved projecting stereoscopic content from the third floor window of a building onto the facade of another building across a busy street at a distance of over 20 meters. The third column in Figure 8 (top) shows the setup used for the second experiment. The projector and cameras remained the same as in experiment #1, and were used in a similar fashion for all the steps in the pipeline.

8.2.3 Procedure. Similarly to experiment #1 the participants were provided with anaglyph 3D glasses. Two stereoscopic videos [(a) a butterfly flying, (b) a roller coaster] were projected first without, and later with the adjustment. The participants were again allowed to move freely in the general vicinity in order to observe from different viewpoints.

8.2.4 Results. The analysis of the responses of the second experiment indicates a significant improvement in the color and depth perception of the viewers. The first and second columns in Figure 8 show a sample frame from the butterfly video, original and adjusted, and projected onto the building facade. Figure 9 shows the statistical significance between the ratings of the participant's perception before and after the experiment for the butterfly video. 83% of the participants who initially rated their perception of 3D with a '3' had increased their rating to '4' when viewing the adjusted video which is higher than the average for '4'. Similar reported improvements can be noted for those who initially rated their perception with '1'. Almost identical results were reported for the second video with the roller coaster. There was an 8% decrease between the two categories for rating of '3' which we believe was due to the responses of the participants of the first night's experiment. Because of the drizzle, the facade had an even darker color than the one captured and used to calculate the adjusted video. In the categories of 'Rate your perception of color and depth', 'Have you experienced nausea, dizziness, or eye strain?', and 'Were you able to perceive 3D?', there were no significant differences from the earlier experiment.

9 CONCLUSION AND FUTURE WORK

Providing low cost minimally intrusive immersive 3D experiences in outdoor public places will open up interesting opportunities for creative artworks by media artists, commercial advertisements, messaging for masses and public education. Anaglyph 3D is clearly a contending technology for this, provided it enables high quality 3D experience. Our work reported in this paper is one step towards improving the quality of 3D stereoscopic content projected onto long range outdoor surfaces such as building facades for public experience. We presented a complete framework to transform stereoscopic anaglyph 3D content so that the quality of depth and colour perception is maintained even after projection onto relatively less accommodating surfaces, i.e., surfaces with complex geometry, texture, color and material properties. We use two fixed long range cameras, one long range projector and one roving video camera to capture the surface images from different viewpoints. With this captured data, we use structure-from-motion to overcome the previously persistent problem of long range calibration. Then through structured light scanning techniques we simultaneously recover shape and reflectance properties of the surface using non-linear optimization. In a final step we iteratively adjust the projection content to mitigate depth and colour perception problems due to the projection surface. We have successfully tested our framework on a number of long range indoor and outdoor scenarios. For validation of the improvement in quality of 3D experience provided by the use of our framework, we specifically set up two user study experiments with participation from random people present at the experiment venues.

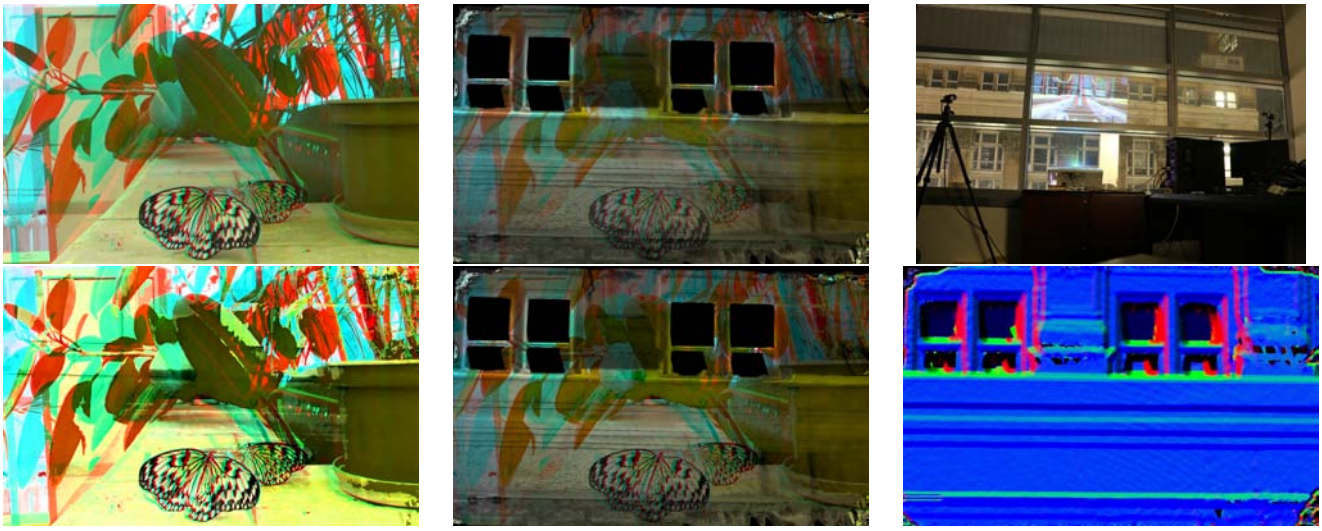

Figure 8: An example frame (from the video) used in Experiment #2. First column: (top) original frame, (bottom) adjusted frame. Second column: (top) projection of original frame on scene, (bottom) projection of adjusted frame on scene. Third column: (top) the projection onto the building's facade; note the excessive ambient lighting present in the scene; distance over 20m, (bottom) the normal map of the projection surface.

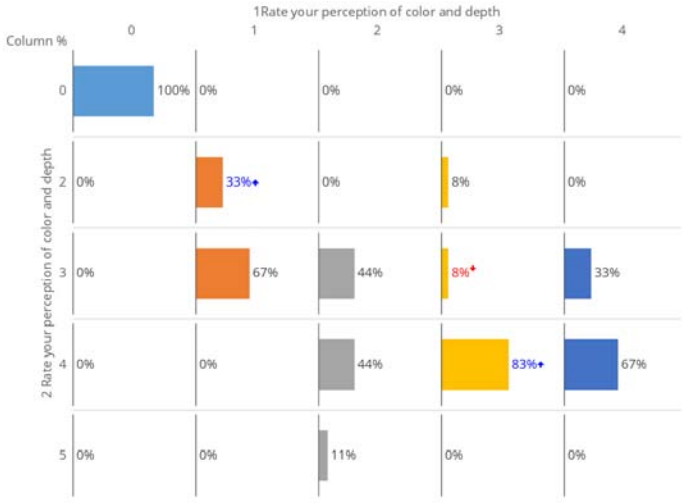

Figure 9: Experiment #2 [Butterfly video]: Statistical significance of participants' questionnaire responses between before (horizontal) and after (vertical) the application of the proposed approach. A likert scale was used ranging from [1,5]. 95% confidence level. Sample size: 37.

Building facades often include glass windows, which scatter light in different directions, letting some of the incident light to pass through. Our present system places cameras in front of the building and hence can only capture reflected light. Correspondingly our framework can only model reflectance properties of the projection surface. We would like to extend our system to model more general light scattering behaviour. We would like to make various computations in the frame work more efficient. One thought which we have is that since the physical nature of the projection surface implies material property coherence, we could consider clustering of the captured pixels/point cloud into regions, and then use this coherence of material property within a cluster to improve computational efficiency. Lastly. we would also want to consider use of multiple cores and GPU computing for improving computational speeds.

10 ACKNOWLEDGEMENT

This research is based upon work supported by the Social Sciences and Humanities Research Council of Canada under Grant No. SO1936, NSERC DGI under Grant No. N00865, Concordia University's Faculty of Engineering and Computer Science under Grant No. VH0003, and the Concordia University CASA Research Grant No. CS1136.

REFERENCES

[1] Bilal Ahmed, Jong Hun Lee, Yong Yi Lee, and Kwan H Lee. 2016. Mimicking an Object Using Multiple Projectors. In *Mixed and Augmented Reality (ISMAR-Adjunct), 2016 IEEE International Symposium on*. IEEE, 61–63.

[2] Daniel G Aliaga, Yu Hong Yeung, Alvin Law, Behzad Sajadi, and Aditi Majumder. 2012. Fast high-resolution appearance editing using superimposed projections. *ACM Transactions on Graphics (TOG)* 31, 2 (2012), 13.

[3] Yatong An, Tyler Bell, Beiwen Li, Jing Xu, and Song Zhang. 2016. Method for large-range structured light system calibration. *Applied Optics* 55, 33 (2016), 9563–9572.

[4] Tyler Bell, Jing Xu, and Song Zhang. 2016. Method for out-of-focus camera calibration. *Applied optics* 55, 9 (2016), 2346–2352.

[5] Oliver Bimber, Gordon Wetzstein, Andreas Emmerling, and Christian Nitschke. 2005. Enabling view-dependent stereoscopic projection in real environments. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 14–23.

[6] Panagiotis-Alexandros Bokaris, Michèle Gouiffès, Christian Jacquemin, and Jean-Marc Chomaz. 2014. Photometric compensation to dynamic surfaces in a projector-camera system. In *European Conference on Computer Vision*. Springer, 283–296.

[7] Panagiotis-Alexandros Bokaris, Michèle Gouiffès, Christian Jacquemin, Jean-Marc Chomaz, and Alain Trémeau. 2015. One-frame delay for dynamic photometric compensation in a projector-camera system. In *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2675–2679.

[8] Maged N. Kamel Boulos and Larry R. Robinson. 2009. Web GIS in practice VII: stereoscopic 3-D solutions for online maps and virtual globes. *International Journal of Health Geographics* 8, 1 (2009), 59.

[9] G Falcao, Natalia Hurtos, J Massich, and D Fofi. 2009. Projector-camera calibration toolbox. *Erasumus Mundus Masters in Vision and Robotics* (2009).

[10] Michael D Grossberg, Harish Peri, Shree K Nayar, and Peter N Belhumeur. 2004. Making one object look like another: Controlling appearance using a projector-camera system. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, Vol. 1. IEEE, I–I.

[11] Qing Gu, Kyriakos Herakleous, and Charalambos Poullis. 2014. 3DUNDERWORLD-SLS: An Open-Source Structured-Light Scanning System for Rapid Geometry Acquisition. *arXiv preprint arXiv:1406.6595* (2014).

[12] Eric PF Lafortune, Sing-Choong Foo, Kenneth E Torrance, and Donald P Greenberg. 1997. Non-linear approximation of reflectance functions. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 117–126.

[13] Alvin J Law, Daniel G Aliaga, Behzad Sajadi, Aditi Majumder, and Zygmunt Pizlo. 2011. Perceptually based appearance modification for compliant appearance editing. In *Computer Graphics Forum*, Vol. 30. Wiley Online Library, 2288–2300.

[14] Behnam Maneshgar, Leila Sujir, Sudhir P. Mudur, and Charalambos Poullis. 2017. A Long-range Vision System for Projection Mapping of Stereoscopic Content in Outdoor Areas. In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. 290–297.

[15] mars.nasa.gov. [n. d.]. Mars 3D Images. ([n. d.]). http://mars.jpl.nasa.gov/mars3d/

[16] Daniel Moreno and Gabriel Taubin. 2012. Simple, accurate, and robust projector-camera calibration. In *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on*. IEEE, 464–471.

[17] Matt Pharr, Wenzel Jakob, and Greg Humphreys. 2016. *Physically based rendering: From theory to implementation*. Morgan Kaufmann.

[18] Gonzalo M. Rojas, Marcelo GÃąlvez, Natan Vega Potler, R. Cameron Craddock, Daniel S. Margulies, F. Xavier Castellanos, and Michael P. Milham. 2014. Stereoscopic three-dimensional visualization applied to multimodal brain images: clinical applications and a functional connectivity atlas. *Frontiers in Neuroscience* 8 (2014), 328.

[19] Christian Siegl, Matteo Colaianni, Lucas Thies, Justus Thies, Michael Zollhöfer, Shahram Izadi, Marc Stamminger, and Frank Bauer. 2015. Real-time pixel luminance optimization for dynamic multi-projection mapping. *ACM Transactions on Graphics (TOG)* 34, 6 (2015), 237.

[20] Roger Tsai. 1987. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal on Robotics and Automation* 3, 4 (1987), 323–344.

[21] Wikipedia. 2017. Anaglyph_3D. (2017). https://en.wikipedia.org/wiki/Anaglyph_3D

[22] Zhengyou Zhang. 2000. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence* 22, 11 (2000), 1330–1334.

[23] Yi Zhou, Shuangjiu Xiao, Ning Tang, Zhiyong Wei, and Xu Chen. 2016. Pmomo: Projection Mapping on Movable 3D Object. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 781–790.