EMBEDDING OF 3D OBJECTS INTO A VIDEO TO DEMONSTRATE ENGINEERING PROJECTS, USING MATCH MOVING TECHNOLOGY

HÄMEEN AMMATTIKORKEAKOULU
HÄME UNIVERSITY OF APPLIED SCIENCES

Bachelor's thesis

Mechanical Engineering and Production Technology, Riihimäki

Spring semester, 2021

Mark Mindubaev

HÄMEEN AMMATTIKORKEAKOULU
HÄME UNIVERSITY OF APPLIED SCIENCES

Mechanical Engineering and Production Technology

Abstract

Author Mark Mindubaev

Year 2021

Subject Embedding 3D objects to a video
to demonstrate engineering projects, using match moving technology

Supervisors Antti Laakso

The aim of the project was to show how a technology of embedding 3D objects into video used in the film industry, namely Match moving, can be used to demonstrate various engineering projects. As an example on how the technology could be used, the author used the Crown Bridges Kruunusillat Project, carried out by WSP Finland and Knight Architects. Based on this, a section of the bridge that was to be recreated was selected, using shots which were taken by the author with a semi-professional drone and with knowledge of the Adobe After Effects program. As a result, the final video was obtained and conclusions were made.

Keywords match moving, visual effects, tracking, virtual reality, 3D

Pages 27 pages

Contents

List of abbreviations

AVC	Advanced Video Coding
CGI	Computer-generated imagery
C4D	Cinema 4D
GrabCAD	Platform with free CAD models
HEVC	High Efficiency Video Coding
ITU-T	International Telecommunications Union's Telecommunications Standardization Sector
LOG	Gamma curve, found on some digital video camera that gives a wide dynamic and tonal range
MPEG	Moving Picture Experts Group
Open GL	Open Graphics Library
OBJ	File format is a simple data-format that represents 3D geometry alone

REC. 709	Standard image encoding and signal characteristics of high definition television
UHD	Ultra-high-definition television
VFX	Visual effects
4K	3840 × 2160

Images and Figures

1 Introduction

The aim of this project was to study the possibility of using modern tracking and motion capture technologies in post-processing and in the introduction of three-dimensional graphics in video by the example of creating a video clip.

In the modern film industry, it is rare that any work is complete without the use of the digital technology. As a rule, the symbiosis of classical tasks, such as directing, camerawork, editing, working with sound, and computer processing is presented. It gives excellent functionality such as: digital color correction and editing can be performed, complete change of the background, adding three-dimensional characters or objects and the usage of full range of available visual effects. Such work can either reduce the project budget, or easily increase it. But if the goal is to create something that cannot be physically done with your hands or does not exist in the real world, then the costs are justified

This thesis analyses the complexity of the process. A beautiful film, a good commercial, or a music video is always based on the classic canons of shooting, in which the camerawork is very important. Rarely, during the entire film, the camera can stand still, in most cases it does not simply rotate at one point on three axes, but moves in space as well. To accurately fit an object into the frame, one needs to resort to tracking technologies, no matter whether it is two-or three-dimensional. The selection of the method should be made on the basis of the goal.

The author of the thesis conducted this project using the Crown Bridges Kruunusillat Project, carried out by WSP Finland and Knight Architects. For commercials and demonstrations of their project, Knight Architects used professional CGI and complete reality recreation. The goal of the present project was to make a similar video using the author's personal knowledge and equipment, and as a result, to analyze how effectively the tracking and match moving technology would work in engineering projects. The author set additional tasks for this project to:

- review the existing technologies,
- analyze available software,

- make post-production in a low-budget video studio,
- find out the pros and cons of the technology.

2 History of tracking and theoretical analysis

The history of tracking technology development has already spanned for more than two decades. It was originally created for military purposes but now is available to be used by ordinary people for peaceful purposes. During this time, it has developed enough to cover a wide range of tasks (Noonan et al., 2009).

Today for each type of problem, there is a specific solution: starting from the simplest task, like camera stabilization when it is standing on a tripod, where position tracking of just one point needed, ending with tracking hundreds of points at the same time, and recently even in the real time. According to NY Castings (2014) tracking is mainly used in cinema but nowadays due to the fact that the technology does not require a high degree of equipment it can be used by any specialist in programs such as Adobe After Effects, Cinema 4D, Nuke, Blender and others. As a result the area of its usage has expanded and now it can be used in engineering, medicine and other fields of science.

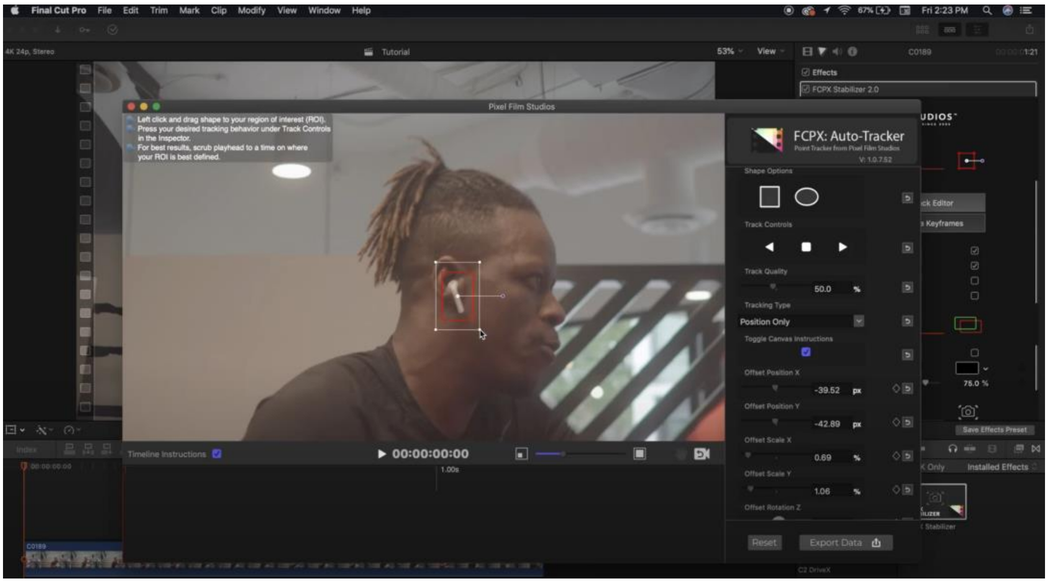
2.1 Overview of project

On this paper several analogues of similar formats are analyzed. Most of them were commercial and social advertising, as well as short films. The author found it relevant to gain knowledge and understanding of the tracking methods used and how these technologies could help during this project. The duration of the video, the features of color correction, the changes in plans, the features of filming, and the method of presenting the main idea were all evaluated here.

2.1.1 Beats by Dr. Dre Commercial

The first analyzed example was a commercial for Beats headphones. The essence of the video was that in each frame, regardless of how the character moves, the headphones always remain in the center of the frame. Tracking of the headphones takes place in the x and y coordinate system. According to Daniel Shiffer’s YouTube tutorial (2019) on creating the same effects, the certain area of the frame is defined, in our case, the headphones, around which a tracking point is created in the shape of a rectangle, as performed in Figure 1.

Figure 1: Tracking box example (Shiffer, Locked-on stabilization effect tutorial, 2019)



The tracker tracks the x and y coordinates of the rectangular area in each frame. As a result of the first stage of tracking, we got a list of coordinates that were later used to stabilize the footage. But what kind of stabilization? Here it became more interesting, because in this case, all footage was stabilized only according to the position of the rectangle while in normal stabilization motion detection methods work completely differently.

As was defined by Argunov school (2018) generally methods used in stabilization work similar to those used in MPEG-like compression. The footage is divided into blocks, for each of them program searches for the most similar block in the previous frame and the offset which is relative to it. The average characteristics for the entire footage are determined from the constructed displacement map. Typically, there are two to four values: horizontal and vertical displacement, and often rotation and zooming. (Although the zoom control does not shake in real shooting, but it is still more convenient to approximate the displacement map by four rather than three values). Then the stabilization module itself comes into action, it builds the optimal camera trajectory by smoothing the existing "chaotic" one, performs displacement, rotation and scaling of frames by the appropriate values and fills the edges. The process of choosing the optimal smoothed path is similar to applying a low-pass filter to the " signal " formed by the displacement vectors. In fact, vibrations with frequencies higher than the cutoff frequency which was defined by a user will be removed. While in this case program makes sure that the coordinates of the area in each subsequent frame correspond to the coordinates in the first frame, which as a result gives the same position of the earphone throughout all frames, regardless of the movement of the character.

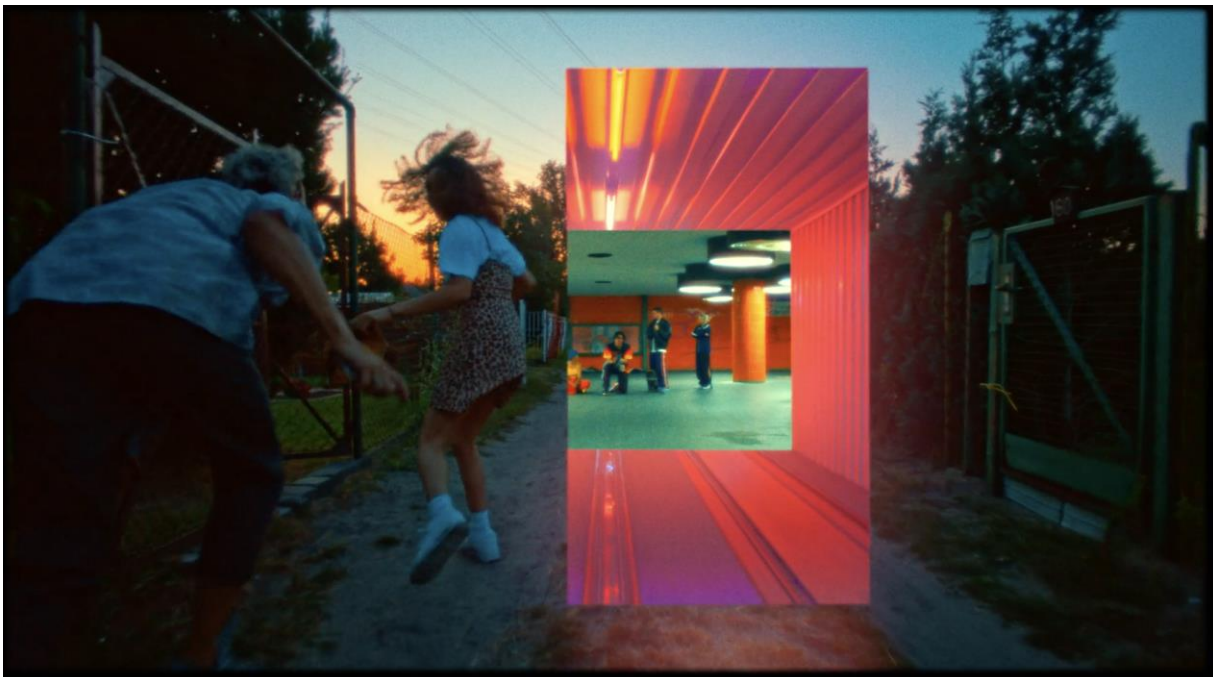
It is important to understand that since tracking was used in our case, any motion could affect our results. But camera work, suitable shutter speed and external stabilizer helped to solve this problem and kept headphones unblurred. More information about motion blur and shutter speed is given in the third chapter.

2.1.2 Nike X End commercial

The famous Nike video commercial in collaboration with the online retail store END was chosen as the second example. Information on the technology used there was not available in full, but after a proper analysis of this example it was found out that this video was shot with planar tracking technology.

As shown in Figure 2 (2020) in this video, the transition from scene to scene went through a certain plane located at the end of each frame, thereby creating a seamless effect. In this case, the plane was a kind of rectangular mask. Since each frame was zoomed and moved forward, naturally the mask/plane could not remain in the same shape, size and location, otherwise a distortion would have occurred and seriously affected the results. Since the plane/mask had stay in the same place and position at all times, tracking came to the rescue, namely planar tracking.

Figure 2: Planar tracking example (Nike x END Clothing Commercial, 2020)



Planar tracking allows collecting such data as: position, scale, rotation, shift, and perspective, as well as to output the resulting track data into the traditional format for tracking corner points. The difference is that flat surfaces are tracked as a whole, rather than individual points. With normal point tracking, you need to determine the location of some objects. These objects must remain unchanged throughout the frame. This task is difficult to manage, since it is often necessary to track frames that were used without taking into account the fact that tracking will be performed in the future (Argunov school, 2018).

In the chosen case, tracking of movement, rotation, and scaling was performed based on the analysis of the movement of a certain user-defined surface. The role of such surface in this commercial was played by a mask. Since the mask remained in the same position, in virtue of the tracking, the desired effect was achieved; it looked unusual, stylish and as seamless as possible.

3 Types of tracking and their differences

The present chapter of this thesis analyses the principles of tracking and the types that should be chosen depending on specific situations. For an initial understanding imagine a random video shot, and select a point or object on it. As a 2D environment it is located in x and y coordinate system, and as a result the chosen point will also be changing according to this system from frame to frame.

For the purpose of a more clear understanding, the author's shot with Volvo, which was performed from a drone during the filming of a commercial for the company PAROC, was taken as illustrated by Figure 3. A specific point to which a 3D object or a plane would be bound or embedded, was selected. By choosing a certain point, any tracking program will provide two windows, the first window, the smaller one - this is what the software is looking for and a second window, a bigger one, is where it is searching.

Figure 3: License plate point tracking in Adobe After Effects example

Let us examine how the tracker works. A tracker sees a certain pattern that was indicated in a smaller box moving from frame to frame and searches for the same pattern which was set in the specified area (a larger box). One of the hardest things during this process is the selection of the correct pattern - the object that the program will search for. In this example it was a car license plate.

Almost all programs work taking into account characteristics of brightness and contrast. In addition, the pattern should remain as clear as possible throughout the entire frame because any mismatch can cause a tracker error. When it comes to clarity, the question immediately arises about the shutter speed and the technical aspects of shooting as seen in Figure 4. If a low shutter speed is used while filming, motion blur will occur, making the picture blurry and making it almost impossible to correctly connect the tracker to some pattern in the frame, thus the final result will be unsatisfactory.

Figure 4: Shutter speed (Tronhouse, 2020)

When choosing a tracking object, corner points, such as a windshield, a road sign, or any object with a flat outline, tracking will perform well in most situations. However, it is not always easy to predict the shot or how it will appear during the post-production stage. In such cases, special mounted markers are used in the professional film industry, which are usually crosses of the contrasting color that are glued to the target piece, ensures that the pattern is accurately chosen; this will simplify obtaining information about the object's movement later on post-production stage, this is shown in Figure 5.

Figure 5: Tracking markers in film industry (Blender3D.ua, 2015)

3.1 2.5D Tracking or Planar tracking

It is not always possible to perform tracking by one, two, or even four points, because the object or plane moves not only in the x and y coordinate systems, but also in the perspective, thereby changing its shape, which causes a parallax effect - that means points that are located closer will move faster than points that are more distant.

Planar tracking provides the same track data as point tracking, including position, scale, rotation, shift, and perspective, as well as the ability to output the resulting track data in the conventional format for tracking corner points. The distinction is that flat surfaces, rather than individual points, are tracked as a whole. Surfaces provide more accurate tracking of coordinates, rotation, and scaling than any point tracking method. Even if the object leaves the scene, there will still be enough data to perform planar tracking, since all points on the surface are involved in the tracking process.

For better understanding of how planar tracking works a study has been conducted in the program/plugin integrated into After Effects - Mocha. Up to date, this is one of the best tracking programs available, well outperforming even the Nuke software.

Planar tracking employs the same algorithms as point tracking. Only a plane is defined as a pattern, rather than points, and the program attempts to locate this plane in the frame. In this case, the trackers will lay the plane while accounting for the change in perspective or turn, attempting to deform the shape to gain the original form exactly.

One of the most important stages is to identify and define the surfaces that will be used for tracking. Often the choice of such a surface is obvious; other times, the target must be broken into many different surfaces and tracking/rotoscoping of a set of surfaces must be performed. According to Lee Lanier (2017) when setting a surface, a spline is drawn around it. The spline does not need to be precise in this situation since the planar tracking algorithm will identify and separate pixels that travel along with the main group of pixels on a given surface from the rest of the pixels.

As it was stated in Mocha 5 essential training (Lee Lainer, 2017) Mocha supports two kinds of splines: X-splines and Bezier splines.

Splines are two-dimensional abstract structures that are entirely self-contained that can be used to construct more complex three-dimensional bodies. Splines are externally a variety of lines, with the form of the line determined by the type of vertices through which it travels. Splines may be plain geometric forms such as rectangles, stars, and ellipses, as well as complicated polylines or curves, as well as text symbol contours. Pierre Bezier, a mathematician, well known for inventing Bezier curves. Renault used Bezier curves and surfaces to computer-design the shape of car bodies in the 1960s. They are currently commonly used in computer graphics (CPU3D.com, n.d.)

Bezier curves are parametrically defined as follows:

Figure 6: Bezier spline formula (CPU3D.com, n.d.)

$$x = P_x(t), \quad y = P_y(t).$$

The t value is a parameter that represents the coordinates of a single point on the line.

The Bezier polynomials for Px and Py are as follows:

Figure 7: Px and Py definition formulas (CPU3D.com, What are splines?)

$$P_x(t) = \sum_{i=0}^{m} C_m^i t^i (1-t)^{m-i} x_i, \quad P_y(t) = \sum_{i=0}^{m} C_m^i t^i (1-t)^{m-i} y_i.$$

Where C_m^i is the combination of *m* by *i*, and x_i, y_i are the coordinates of the points of the landmarks P_i. The value of *m* can be considered both as the degree of a polynomial, and as a value that is one less than the number of reference points.

In general, X-splines are best suited for tracking, particularly when the perspective changes. Bezier splines are more general, and they are used in the majority of software.

In Mocha, tracking data can also be used for stabilization; the algorithm uses the data collected as a result of tracking to correct for frame movement. (Boris FX, 2018)

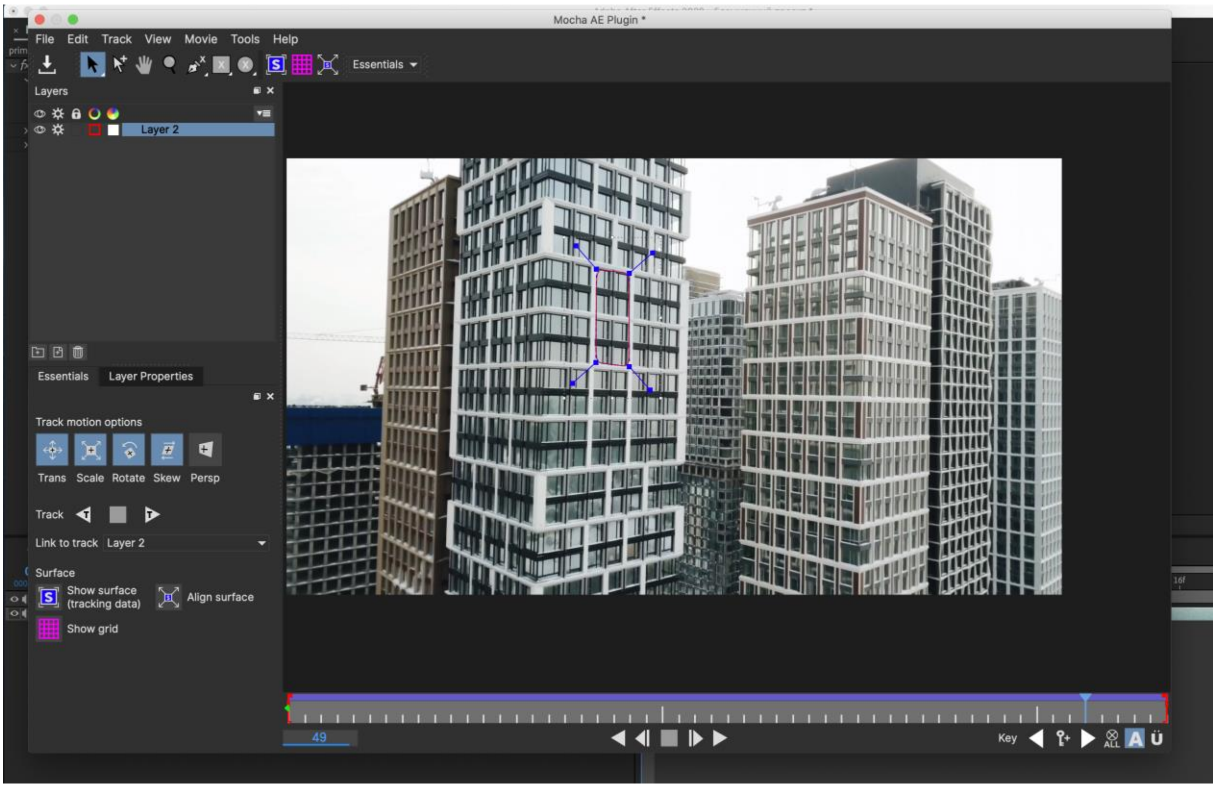
Stabilization is used to:

- eliminate excess distortion or camera shake,
- mask the movement of the layer to make future work with it easier; the movement is then applied to the final composition.

Furthermore, unlike point tracking, a search zone in planar tracking doesn't exist since Mocha attempts to find a pattern throughout whole picture. The only restriction that must be imposed is that one's software should not be provided with needless detail, it must state intentions specifically. Following the selection of the desired plane, the software allows to specify which data the tracker will monitor: Transform, Scale, Rotation, Shear, and Perspective (TSRShP in short), where Shear is the shift.

My drone video footage seen in Figure 8 from a commercial shot for the PrimePark residential complex in Moscow was used for the experiment of planar tracking in Mocha AE.

Figure 8: Planar tracking in Mocha AE example, 2021



As a part of the experiment, the benefits of using planar tracking were discovered, and the principles of work were comprehended. These benefits included:

- no need for extra clip processing (for example, color correction),
- no need for manual monitoring, such as when the track point leaves the video.

3.2 3D Tracking and Match moving

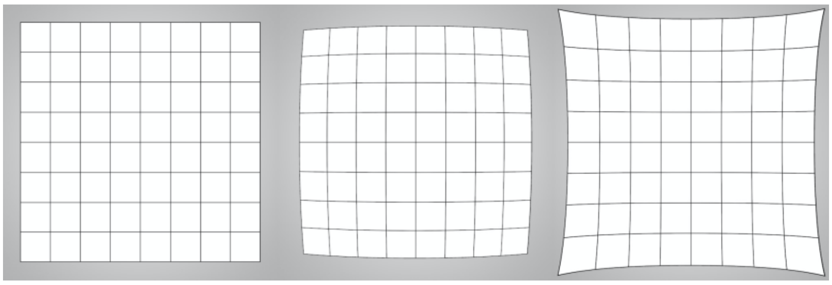
Working with 2D artifacts and their tracking in the standard x and y coordinate system is not always sufficient in film production. In these cases, 3D tracking, Motion tracking, and Match moving come to the rescue.

When operating in 3D space, a coordinate grid with a coordinate center that corresponds to 0,0,0 is needed. In 3D tracking, the z coordinates are added in addition to the x and y coordinates, which define the distance, that is, how far the target is in the frame. 3D tracking, in a nutshell, is used to map dynamic camera motions. The aim is to build a simulated space that matches the video footage completely (or partially). Owning this, positioning of 3D objects possible anywhere in the frame. When making this type of tracking, the most crucial thing is to provide enough points to create a 3D scene. The tracking points should be placed in both the foreground and background of the video. The depth of a 3D scene is determined by dividing the speed of movement of objects in the foreground by the speed of movement of objects in the background. The more objects needed to be tracked on both plans, the more accurate the tracking will be. In 3D space, one of the most important things is to build a virtual camera that corresponds to a real camera that exists in the same coordinate systems as our scene. In a 3D scene, the camera can be applied in three parameters: translation, rotation, and scale. (Argunov School, Nuke compositing, 2018).

When the camera is tracked, new element can be placed among other objects in the video, simulating its location in the scene. After tracking the movement of an object, transformation of the movement can be performed (or calculation of the derived movement based on it) to a new element. But not always the movement in the camera can be solved as correctly as possible, since an important factor is the distortion, which must be seriously taken into account.

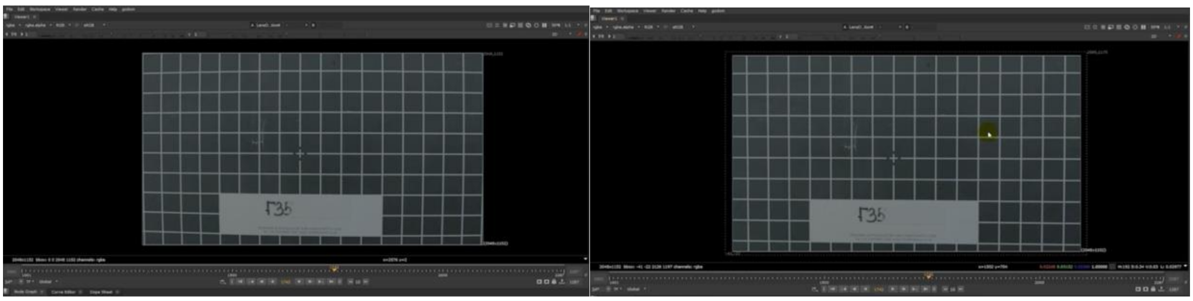
Distortion is a form of optical aberration, or in other words , image distortion. Distortion is a geometric aberration, which means that it alters the image's proper geometry (Learn Foundry, n.d.). Distortion can be classified into two types: barrel-shaped (convex distortion) and pillow-shaped (concave distortion). See Figure 9, Distortion types.

Figure 9: Distortion types (Learn foundry - Advanced Compositing with NukeX and Nuke Studio, n.d.)



In the case when the distortion is present and acts in such a manner that objects closest to the camera begin to shift inappropriately, movement of the virtual camera can be inaccurate, resulting in incorrect tracking and snapping of the target. To correct or avoid lens distortion, firstly, lens distort should be obtained and then an undistortion must be performed. (Figure 10, 2018). In such cases, the distortion grid, which comes in two varieties, chessboard and grid, is used. Most 3D tracking programs can automatically analyze the grid that was provided.

Figure 10: Lens Distortion fix by grid in Nuke (Learn foundry - Advanced Compositing with NukeX and Nuke Studio. Estimating Lens Distortion Using a Grid, n.d.)



In cases of linear motion of the camera, lens distortion may not affect tracking significantly, but if the movement is more complex, the tracker may begin to behave completely inadequately.

Since proper orientation of the scene is the essential part of 3D tracking, the plane of the surface should be exactly at 0. This is accomplished by constructing a plane at zero, i.e., constructing a so-called floor for the scene. When the ground plane is set, the tracking points move with the object and do not interfere with previous tracking. In addition, when tracking and integrating of 3D objects is performed, the size of the whole scene is critical, since if it is not set, the virtual camera and program itself would not be able to determine what scale the actual scene was. Perhaps it was shot somewhere in a large area, or vice versa, on the macro lens in the sandbox. Technically, if the distance between any two points a and b in real-time shooting is known, it can be put safely into the software and it will accurately map the scene.

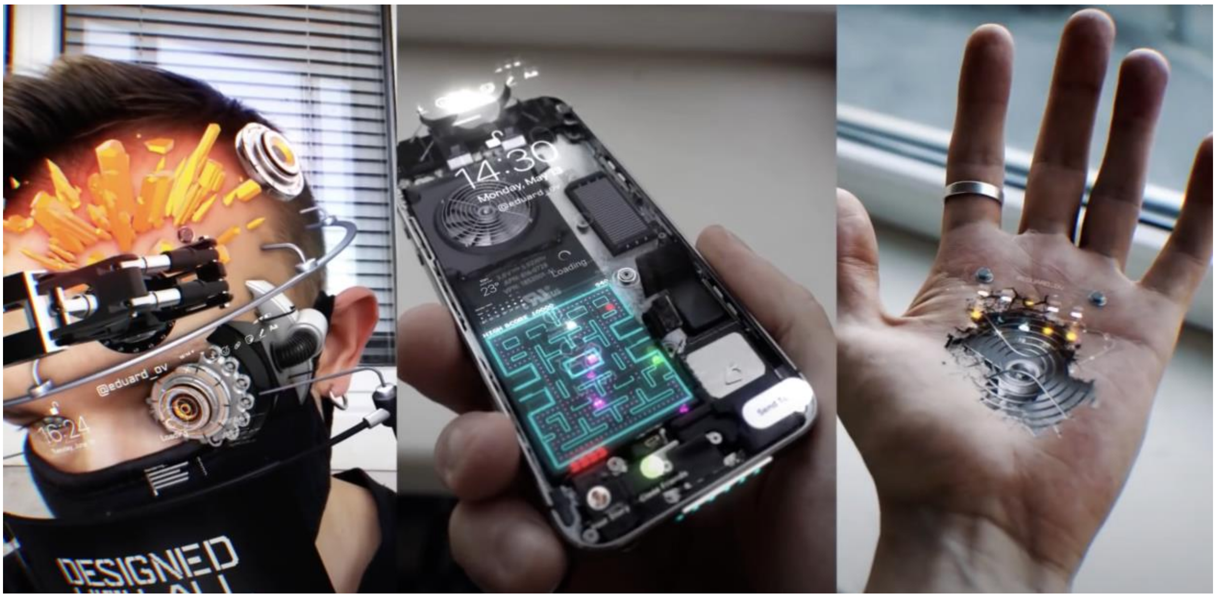
Match moving process takes place in three major steps as follows:

- Tracking, as a result of which set of trajectories of the movement of pixels corresponding to the movement of objects in the frame is received. It is necessary to track objects that were stationary in the original video sequence (in real life) — static props, buildings, trees and etc.

- The next stage is Setup - the process of establishing logical relationships between the tracked points. As a result of this step, the program will be able to establish a correspondence between a point (pixel) in a flat video sequence and a point in a virtual (three-dimensional) space.

- Final stage — Solving and Fine-Tuning . At this point, the program performs calculations and sets the approximate location for the points from the previous step. Based on the location of the points in space, the trajectory of the camera is calculated. During the setup process, the data that helps the program to determine nott only the trajectory of the camera, but also the orientation and scale of this trajectory in space is entered by user additionally. The program does not know whether the camera moved vertically or horizontally. By "tying" the tracked points to specific positions in space, the user sets the scale and orientation of the camera. Usually, parts 2 and 3 are repeated until the result is satisfactory.

Figure 11: Match moving in After Effects (Eduard Mykhailov Mad VFX in AE course, 2019)



4 Visual effects in Engineering

VFX technologies, high-quality rendering, and other digital effects are fully used in contemporary engineering and architecture, helping to mimic life or function on something that does not yet exist in the physical world. However, most of the time, CGI is a massive task that requires a huge amount of commitment, time, and, of course, resources. (EDUCBA, Difference between VFX and CGI, n.d.). Sometimes, engineers have completed all of the calculations, architects have designed the model, but no one will see an indication of how it would appear in comparison with the actual world, because recreating reality is time-consuming and costly. In those circumstances, classical architectural renderers are used.

But why should we replicate reality when 3D tracking and Match moving technology can blend 3D objects with actual reality? This method can help engineers correctly interpret the potential object in relation to the current structures around them, and can as well help architects to correctly visualize and represent unity with the environment. Following that, I have performed the analysis of the application of this technology based on the Crown Bridges Kruunusillat Project, which was designed by WSP Finland and Knight Architects and was awarded by the Helsinki City Council (Figure 12, 2013). In terms of personal operator expertise and low-budget video studio, I made an attempt to reconstruct their project without use of CGI, but only within the context of tracking and match moving technology.

Figure 12: Crown Bridges Kruunusillat Project render (Knight Architects, 2013)

4.1 Advantages and disadvantages of using tracking in engineering

Following an analysis of current tracking types and the implementation of 3D objects in video, it was determined that 3D tracking is ideally suited for solving this problem.

But what issues could occur if this approach would be used? The most critical point is that the world of the video in which the 3D object may be embedded will not be ready for development because, for example, where there may be an exit from the bridge, there is still a lot of trees, so the effect will overlap with reality and not correspond to the future. The quality of the 3D model may also have a significant impact on the outcome, which will influence the visual appearance of the future project. The availability of the location where the clips for match moving can be taken will also be a major issue, since not all places can be reached and not all angles for potential tracking can be shot anywhere, unlike a complete recreation of reality on a computer.

The advantage of the technology is that the implementation of visualization does not require a complete reconstruction of reality, and the number of necessary 3D objects is reduced significantly, thereby reducing the cost of the project. It also makes it possible to see an example of how it will look relative to the current reality much earlier than the start of construction, thereby helping to make the right decisions and not make certain mistakes.

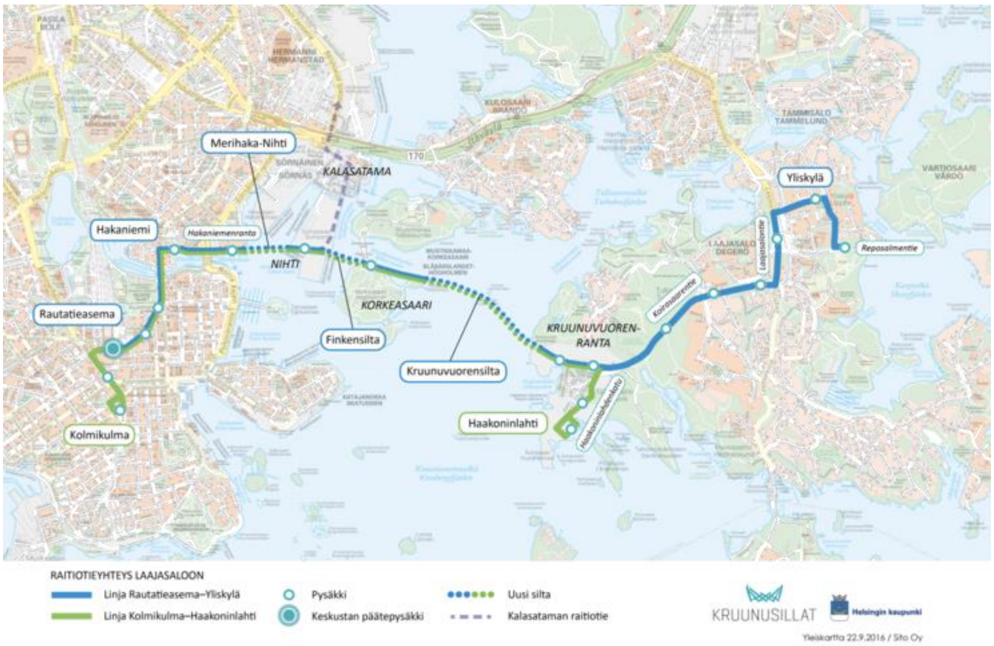
5 Main project and its pre-production

As it was defined the main goal of the project was to show how the technologies of 3D tracking can be used to demonstrate various engineering projects and what are the main differences between completely recreating reality and adding the unreal to real. For the more clear picture, Crown Bridges Kruunusillat Project, carried out by WSP Finland and Knight Architects, was taken. On its basis, within the budget video studio, personal familiarity with camerawork, and knowledge of post-processing in Adobe After Effects, a certain section of the bridge has been chosen, and the technique of integrating 3D objects in video have been tested. Based on the project's findings, it aimed to help determining if Match moving is preferable to CGI.

Before proceeding to any filming, and even more to the post-production stage, it was needed to seriously analyze all the upcoming stages, in order to avoid unnecessary problems on other phases. The first important aspect was to decide which part of the bridge can be recreated and where it would be necessary to shoot from the drone. In the open access network, a map of the construction of future bridges can be found (Figure 13, 2013). After analyzation, the section between Sompsaari and Korkeasaari was chosen, since this is the shortest section connecting the two islands, and in the conditions of shooting from an amateur drone, it would work the best, since all possible and necessary angles would be captured without the risk of losing communication with the drone.

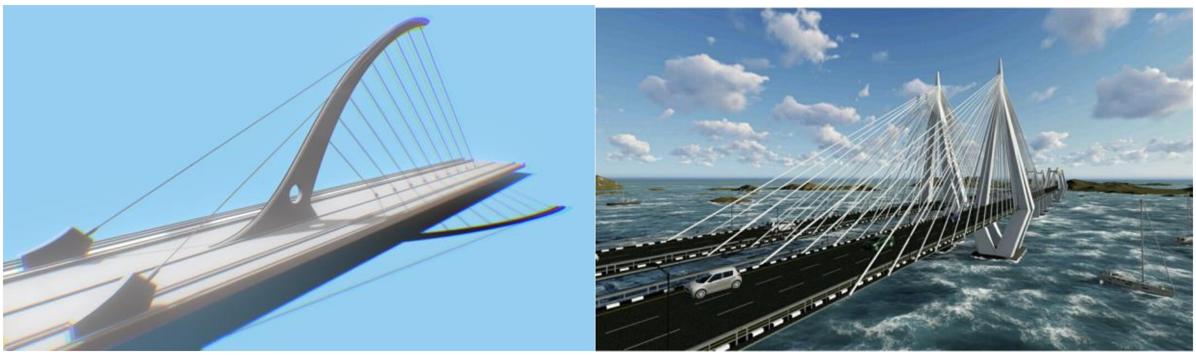
The next key factor was to determine the best possible time to shoot, since in order to track bridge on the post-production phase, a distinction between them needed to be provided, since there is no way to place markers. The sunset time has been chosen as the sun would slowly set and the necessary contrast would be present.

Figure 13: Crown Bridges Kruunusillat Project map (Knight Architects, 2013)



The selection of a comparable 3D model, which has been used in the video series, then took place. This was an important point to consider in order to get a better view of what angles can be used to make the future image appear as stunning as possible. Since my role in this project did not require me to build my own 3D model, one from open and free sources was chosen. Following an analysis of the available 3D bridge models on the GrabCAD resource, two bridge options (Figure 14, GrabCAD) as best suiting the implementation in the scene.

Figure 14: Selected 3D bridge models (GrabCAD n.d.)



5.1 Drone shooting

A convenient shooting day on the island of Sompasaari was chosen. For the filming, DJI Mavic Pro 2 drone was used. This drone can film in 4K resolution (3840x2160) at a frame rate of 30 frames per second. At the starting moment of the project, the light was ideal, but there was a problem: there was strong wind, which could potentially interfere with the stabilization when shooting. The drone was launched into the air, but after a short period, it was decided to reschedule the shoot for a less windy day, because the chance of dropping the drone and receiving low-quality footage was very high. After a while, on a less windy day the necessary material was shot.

The key challenge was to shoot the most beneficial angles, where camera movement would be most profitable when making a virtual camera and its movement in post-production. The bridge had to be represented and shot with the assumption that the object was present in the frame, which was an interesting challenge in shooting. The camera movements were chosen as follows: circular flyovers clockwise and counterclockwise, straight flyovers forward and backward, and long-range general plans with minimal camera movement.

5.1.1 Video format, color space and codecs

The selection of the appropriate format was an essential part of the shooting process, so that there would be no problems later in the post-production stage. The UltraHD format was selected because it is not only a variety of 4K standards, but also the standard of ultra-high definition television. The resolution of the UHD is 3840 x 2160 pixels. This format was chosen because high resolution was needed so that the loss of information when zooming in on our frame would be negligible.

The selection of the best codec was the next critical consideration. Since the used drone was one of the most recent models on the market, it could film in a High-Efficiency Video codec (HEVC), also known as the H. 265 codec, which compresses more than twice as well as the best video codec for Blu-ray. It is the latest counterpart to Advanced Video Coding (AVC), a codec also known as H. 264 and one of the major Blu-ray encoding systems. The concept behind HEVC is to have the same amount of image quality as AVC while improving compression, so that a video file compressed with this codec is half the size. This is important because the smaller the file, the less time it takes to make and the faster and more convenient it is to work with.

In 4K video, the data stream is much better than in HD video. Although most people are still used to the fact that the H. 264 codec has an advantage over MPEG-2, the Motion Picture Experts Group and the International Telecommunications Union's Telecommunications Standardization Sector (ITU-T) have already started work on the next wave of video encoding. For each transition to a different norm, either the video volume doubles at the same resolution or the picture quality doubles at the same volume. This was mostly accomplished as a result of increased use of AVC.

To begin, the latest codec examined multiple frames at once to determine what is not moving in the picture. The vast majority of the video of most scenes of a TV show or movie do not vary much. Consider a scene in which someone is speaking. The majority of the picture is occupied by the subject's head. The background does not change much for many frames. Similarly, the majority of the pixels that comprise a face are unlikely to change significantly, except for the lips, of course. Instead of encoding every pixel from every frame, the first frame is encoded, and then the modifications are encoded after that.

The scale of the region that these modifications are looking at is then increased by HEVC. Larger and smaller "blocks" are needed for increased efficiency. In HEVC, they may be bigger, smaller, or with different shapes than in previous codecs. For example, larger blocks proved to be more effective. On Figure 15 (2014) the left is AVC / H. 264 macro-blocking. As can be seen, on the right there is much more flexibility and larger sizes, for the HEVC / H. 265 encoder.

Figure 15: H.264 and H.265 codecs difference (TrueConf, 2014)

6 Post-production process

The first step was to go over the content and to choose the best and most appropriate takes. The key requirements were image stability (to maximize the minimum use of additional stabilization), a smooth camera rotation, and frames of both banks visible (this was necessary so that during the tracking process our 3D model could grab the tracking point from both sides and make it easier to work).

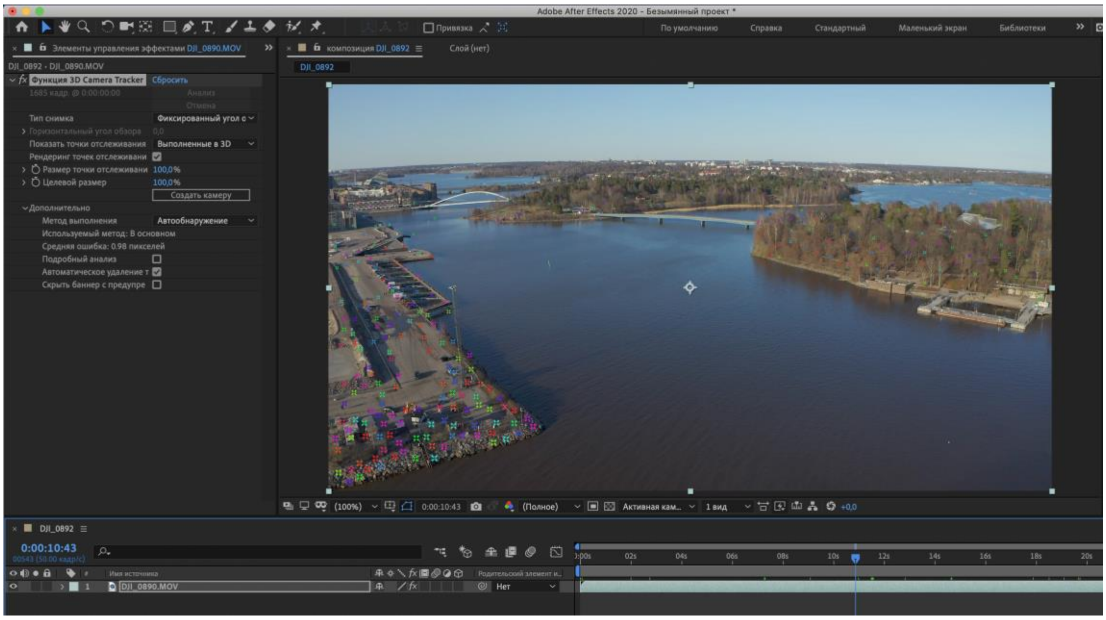
After the material selection, I was able to start working and experimenting, but since I had sufficient knowledge of motion design and 3D, primarily in the Adobe After Effects software, I chose it as the main tool for this project. We had the required 3D tracking tool in After Effects, but not for 3D models. To do so, we were to use the Element3D plugin, which was developed by the glorious Andrew Kramer and his wonderful team from Videocopilot.net.

These professionals created what humanity has been looking forward to since 1993. Starting from the earliest iterations of After Effects, it lacked support for three-dimensional models, despite the fact that all of the prerequisites for this were in place – cameras, lighting, viewports – but for whatever reason, such an obvious thing was achieved only by Kramer and the Company after nearly 20 years. The plugin supports environment mapping and motion blur, depth of field, and can also export rendering paths. The program's key engine is the well-known graphics plugin-Open GL, which is characterized by its wide capability. It is also worth noting that it operates at a high speed, allowing you to render graphics in the shortest amount of time. In terms of format support, Video copilot only supports a few options: C4D, OBJ, and 3DK but this was enough for us.

6.1 Workflow in After Effects

According to Jas Davis's YouTube tutorial (2020) the first step is to import the chosen files into the newly developed project. Thus I chose the shortest and most suitable video fragment for the first test of incorporating a 3D model of the bridge in the video series. The video was enhanced with the 3D Camera Tracker effect. The 3D camera tracker feature analyzes video episodes to provide information about camera movement and three-dimensional scene. After the analysis, the effects of my video sequence determined the entire environment, mapped out the 3D space, and determined where the camera moved in 3D space. As it shown in Figure 16 (2021), a lot of 3D points appeared there.

Figure 16: 3D tracking results in Adobe After Effects (2021)



The next step in the software was to build and define the so-called floor. I have made reference planes so that when I place 3D objects in the frame, they would automatically be directed correctly, and then set the coordinate center of the 3D universe. After creating several reference planes, I was able to start working directly with the 3d object. To accomplish this, a null object has been generated and the Element 3D plugin was mounted on it. I loaded the bridge model and set its orientation to bottom within the plugin, thus connecting our object to the ground plane. The only disadvantage of using Element 3D was that the models were loaded without textures, but they could be applied from inside the plugin.

The model appeared in our video sequence after I had completed the necessary manipulations. My job now was to accurately scale the bridge so that it appeared as realistic as possible in the shot. Once our object was properly placed, the shadow added volume to our scene, which was the important step in creating such a project. The shadow was made by means of creating another ground plane for our object, but material was applied to it, namely, Matte Shadow. This material is fundamentally different from other types of materials, as it is not visualized. The intention of this material is not to show off the surface in any way, but rather to conceal it. When you apply the Matte Shadow material to an object, it creates a kind of hole in the scene, allowing one to cut off the geometry behind it and to show the background. This allows one to simulate the mixture of the background and the items in the scene. Furthermore, objects with the Matte Shadow content appear invisible to the background image, and other objects may cast shadows on them, giving the impression that the objects are casting shadows directly on the background.

The next phase was to add more light to make the subject look good in the light conditions that were present when the video was shot. This was accomplished with the Spot Light tool. A spotlight in After Effects functions just as a real-world spotlight. It is a single point that one can aim around and point at things. They are usually smaller, more circularly oriented lights with power of how wide or narrow the beam is as well as how sharp the falloff is. Spotlights are often used to illuminate a small area of a picture, leaving the remainder in black shadow with a very sharp fall off.

Once the lights were correctly positioned in the scene, I was able to proceed with styling and give the video a single color. I used a new program and so called LUTs to do this.

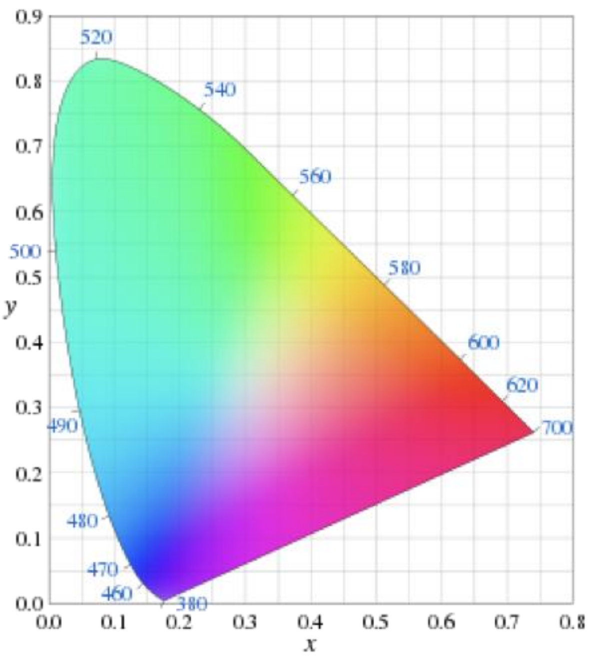
6.2 Colouring and LUTs

In this chapter, the color, color spaces and LUTs are analyzed, as well as how to apply all of these principles to the style of our sequence.

The term look-up table is represented by the abbreviation LUT. This is not the only way to adjust the video's color and basic parameters, but if the others are lossy for the clip, LUT is not. To compare, one can color correct any video using the photographer's basic set of sliders for changing brightness, contrast, and curves, and then do the same using LUT profiles. In the first case, since footage is usually 8-bit, one will only lose tones when modifying and improving the contrast. And if you do the same thing through the LUT, the tones would be lost minimally or not at all.

A picture or video file can be interpreted mathematically not only as a sequence of zeros and ones, but also as a map of the environment with two axes of coordinates as can be seen in Figure 17: vertically and horizontally. It displays the numerical color codes illustrated in Figure 18 for each pixel in the main points. Anyone who has worked with Photoshop knows that binary code will represent any color in the color space specified for the image.

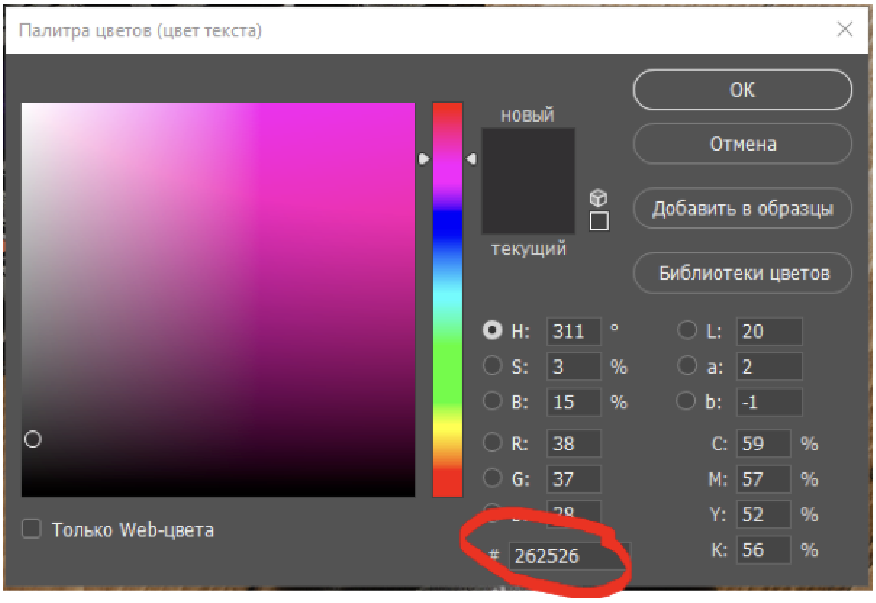
Figure 17: CIE 1931 Colour space diagram (Blatner, Fraser, Real World Photoshop CS, p. 179, 2014)



Space is a coordinate system that determines values for various codes. For instance, a certain code in the sRGB space refers to one color; but, if the same color code is used in the Adobe RGB space, the color would be subtly different. In a nutshell, each pixel corresponds to a specific color code, representing its brightness and color in the corresponding color space (or profile). (Blatner, Fraser, Real World Photoshop CS, p. 179, 2014)

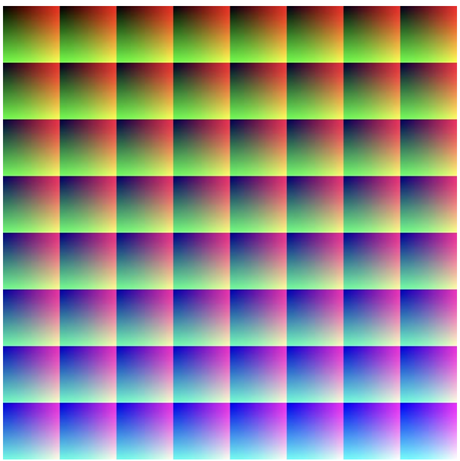
LUTs are classified into two types: technical LUTs and styling LUTs. Technical LUTs are used to convert the camera's color profile to the REC. 709 color space because the camera cannot initially film in it. If one uses a professional camera that records in REC. 709 or LOG, he will not need to use a technical LUT. The second type of LUT is stylizing (creative), which employs a variety of filters such as separate shadows and lights (split-toning), contrast, white balance adjustment, saturation. Playing with LUT allows one to customize the image to his taste.

Figure 18: Colour as a code in Adobe Photoshop



Since I am a photographer, the ExportLut plugin was used to colorize the film. John R. Ellis created this plugin, which helps to make LUTs from Lightroom presets or the develop settings of images and video frames. In other words, this plugin creates a LUT box as seen in Figure 19 based on our Lightroom settings. When the LUT is a picture with the colors arranged in the specified order. In the LUT image template, the colors are mapped to themselves.

Figure 19: LUT box (Kosmonaut Blog, 2017)



The color processing is achieved by exporting any frame from the video sequence in jpeg format and then editing it in Lightroom as a regular picture.

7 Problems and their solutions

During the post-production process, a number of problems were encountered:

1. On the second island, where the final portion of the 3D model of the bridge was situated, there was a shortage of tracking points.

This problem arose due to the fact that the preparations for the shooting did not take into account the fact that there was a high number of trees on the second island in the proposed location of the bridge. Since they had such a similar pattern, the tracker had a difficult time identifying any particular piece. This fact also did not work in my favor because it was impossible to specify the right reference plane due to the wrong location of tracking points and their limited number. The alternative was to choose another island with enough monitoring points for our 3D model to track on.

2. Initially wrong choice of 3D model

In the process of preparing for the post-production stage, the 3D model of the bridge being implemented was incorrectly selected. It turned out to be too big and unsuitable for the style of the video. The necessary textures were also missing. The problem was solved by simply choosing another, more suitable model.

The rest of the other post-production stages went off without problems and were completed as planned.

8 Conclusion

The key aim of this project was to examine how relevant the use of VFX technology, specifically Match Moving, was for the fields of engineering and architecture, using the Crown Bridges Kruunusillat Project, which was performed by WSP Finland and Knight Architects. Commercials that used tracking technologies were examined, as well as the technical implications of various methods of tracking and their applications.

Following that, the decision was made to use 3D tracking as the most appropriate tool for the purposes of my project. Furthermore, different problematic issues that might exist when operating on such a project were considered, as well as the types of software that could be used. The Adobe After Effects package and the Element 3D plugin were chosen due to extensive personal familiarity with these programs.

The next important step was to shoot from a drone, during which different video formats, codecs, and potential bugs were examined. The actual filming took place in an area where a portion of the bridge was to be constructed in the future. No particular problems appeared after shooting in the post-production period, as a large number of materials were studied with regard to the embedding of 3D objects into the video in Adobe After Effects.

Based on the results, it was concluded that the use of VFX and tracking technologies is extremely useful for engineering and architecture, and that it can have a great additional value in project demonstrations while also significantly lowering the cost of presentations and visualizations. This conclusion was reached based on the fact that in the CGI sector and the creation of a completely new reality for projects of such a scale as the Crown Bridges Kruunusillat Project, labor and project remuneration is performed at a very high cost, which greatly exceeds the rates in the field of match moving.

After personally going through all the stages of creating a project to embed 3D objects into a video on a zero-budget basis, I can honestly claim that with more professional equipment and experience in the field of Match moving, the outcome at the lowest cost would be impressive and provide high quality. I consider the findings of my study and project have been successful.

References

Aldinger P. (2020). *NIKE x END CLOTHING // COMMERCIAL* [Video]. Vimeo. https://vimeo.com/364473275

Blender3D (2016). *Getting to grips with Motion Tracking*. Retrieved April 17, 2021, from https://blender3d.com.ua/razbirayemsya-s-motion-tracking/

Boris FX. (Febrauary 13, 2018). *What is Mocha and planar tracking?* [Video]. YouTube. https://www.youtube.com/watch?v=ZXGqI389GvY

Blatner F. (2015). *Real World Photoshop CS*. Retrieved March 11, 2021, from http://canopuscomputing.com.au/gallery2/d/8631-1/Gamuts_comparison-B_F.jpg

CPU3D.com (n.d.). *What are splines?.* Retrieved April 25, 2021, from http://cpu3d.com/lesson/chto-takoe-splayny/

Davis J. (2020, May 5). *3D Object Tracking & Animation in After Effects using ELEMENT 3D + Filming with Osmo Pocket!* [Video]. YouTube. https://www.youtube.com/watch?v=6IHUR4du590

EDUCBA (n.d.). *Difference Between VFX and CGI*. Retrieved March 9, 2021, from https://www.educba.com/vfx-vs-cgi/

Gatt J. (2014). *Motion Capture Actors: Body Movement Tells the Story.* Retrieved April 12, 2021, from WayBack Machine: https://web.archive.org/web/20140703113656/http://www.nycastings.com/dmxreadyv2/blogmanager/v3_blogmanager.asp?post=motioncaptureactors

Godomin I. (2018). *Argunov School Course - VFX ARTIST: NUKE COMPOSITING Motion.* Retrieved April 21, 2021, from: https://argunov.school/courses/vfx-artist-nuke-compositing/

Knight Architects (2013). *Helsinki's international landmark you can see from the sky.* Retrieved April 12, 2021, from https://www.knightarchitects.co.uk/projects/kruunusillat

Kosmonaut Blog (2017). *Color Grading and Correction.* Retrieved March 9, 2021, from https://kosmonautblog.wordpress.com/2017/04/26/color-grading-correction/

Lanier L. (2017). *Mocha 5 Essential training*. Retrieved April 18, 2021, from https://www.lynda.com/mocha-tutorials/Planar-tracking-vs-standard-2D-tracking/601820/644179-4.html

Learn foundry (n.d.). *Advanced Compositing with NukeX and Nuke Studio*. Retrieved April 23, 2021, from https://learn.foundry.com/nuke/11.3/content/comp_environment/lens_distortion/analyzing_distortion_grid.html

Mykhailov E. (2019). *MAD VFX in After Effects*. Retrieved March 6, 2021, from https://motiondesign.school/products/mad-vfx-in-after-effects

Noonan D., Mountney P., Elson D., Darzi A., Yang G-Z. (2009). *A Stereoscopic Fibroscope for Camera Motion and 3D Depth Recovery During Minimally Invasive Surgery.* Retrieved April 12, 2021, from Research Gate: https://www.researchgate.net/publication/221073737_A_stereoscopic_fibroscope_for_camera_motion_and_3D_depth_recovery_during_Minimally_Invasive_Surgery

Shiffer D. (2019, June 10). *LOCKED-ON STABILIZATION EFFECT (Beats By Dre Inspired) | Final Cut Pro X Tutorial* [Video]. YouTube. https://www.youtube.com/watch?v=iPlpAuY7nO4&t

TronHouse. (January 13, 2020). *Shutter speed* [Image]. https://tronhouse.com/part-3-tim-hieu-ve-shutter-speed-toc-do-man-trap-trong-nhiep-anh-p8256

TrueConf (2014). *How 4K differs from UltraHD.* Retrieved April 27, 2021, from https://www.knightarchitects.co.uk/projects/kruunusillat