

Improvements in the visualization of stereoscopic 3D imagery

Luis E. Gurrieri*

Abstract

A pleasant visualization of stereoscopic imagery must take into account factors that may produce eye strain and fatigue. Fortunately, our binocular vision system has embedded mechanisms to perceive depth for extended periods of time without producing eye fatigue; however, stereoscopic imagery may still induce visual discomfort in certain displaying scenarios. An important source of eye fatigue originates in the conflict between vergence eye movement and focusing mechanisms. Today's eye-tracking technology makes possible to know the viewers' gaze direction; hence, 3D imagery can be dynamically corrected based on this information. In this paper, I introduce a method to improve the visualization of stereoscopic imagery on planar displays based on emulating vergence and accommodation mechanisms of binocular human vision. Unlike other methods to improve visual comfort that introduce depth distortions in the stereoscopic visual media this technique aims to produce a gentler and more natural binocular viewing experience without distorting the original depth of the scene.

Luis E. Gurrieri, "Improvements in the visualization of stereoscopic 3D imagery", 3D Research, Volume 6, Issue 25, September 2015.

Online date of publication: Jul. 3, 2015. The final publication is available at Springer via:
<http://dx.doi.org/10.1007/s13319-015-0058-3>

*Luis E. Gurrieri, University of Ottawa. luis.gurrieri@uottawa.ca

1 Introduction

The correct visualization of stereoscopic imagery has to take into account the binocular mechanisms of human vision that prevent eye fatigue. Two processes interact to maintain a binocular vision voided of discomfort: vergence and accommodation. The former is defined as the simultaneous toed-in movement of both eyes to obtain or maintain a singular view of a scenic element [1]; the latter is based on adapting the visual depth-of-field (DOF) or eye focal power to clearly see that scenic element.

Current technologies to produce and display stereoscopic imagery are mature to make outstanding depth effects without producing eye strain. However, in some three-dimensional (3D) displaying scenarios visual discomfort may arise, for instance, in immersive applications or whenever the display is relatively close to the viewer [2]. This problem is mainly due to overusing large DOF shoots, assuming true that superior image quality comes hand-to-hand with sharply rendered images.

In recent years, stereographers have put great attention to prevent excessive disparities at screen distance, which is usually far from the viewers, so vergence-accommodation conflicts are naturally neutralized. Most of the complains regarding eye strain in stereoscopic media come from 3D gaming and immersive applications where the screen is relatively close to the viewers' eyes [3].

The tolerable limits of visual comfort while watching stereoscopic media are well known [4, 5, 6, 7]. These parameters have been tested and recorded as acceptable tolerance for the eye vergence movement for a given distance between the viewer and screen [8]. An object can be rendered to appear virtually on the screen (zero parallax), behind it (positive parallax) or in front of it (negative parallax) by controlling the screen parallax or the amount of horizontal overlapping between left and right projected views seen on the screen. This can be done during acquisition time by controlling the width of the stereoscopic baseline, or by modifying the vergence angle between the optical axes of both cameras in the rig, or by controlling the overlapping between the left and right images projected on the screen.

One problem is the sudden transition between 3D spaces, which leads to rapid changes in vergence eye movement while forcing to maintain focus at screen distance [9, 10]. The eyes strain is exacerbated when the 3D budget¹ is at the limit of the tolerable vergence-accommodation range. It is not uncommon to find reports of headaches and general rejection of 3D technology after have being subject to this *visual gymnastic* over a couple of hours.

Fortunately, there is a simple solution based on currently available *eye tracking* technology. The viewer's gaze direction conveys useful information for rendering customizable stereoscopic views for personalized experience. There is not much that can be done in a shared screen scenario; but the stereoscopic visual media can be streamed to viewers wearing head-mounted displays with the capability to locally customize the displayed views based on where in the scene each viewer gaze direction is [11, 12].

In this paper, I propose a technique to reduce the visual discomfort when 3D content is displayed on planar screens located at a short distance from the eyes. In order to do so, I structured the paper to first introduce the theoretical background that supports the idea, then present the vergence-accommodation method to be applied on any stereoscopic image in a eye-tracking scenario and finally show a few experimental results based on emulating an eye-tracking feedback system. The main goal of this paper is to present the image processing technique used to correct stereoscopic imagery in a navigation setting; neither its goals is about building a testbed nor is about reporting test results in human subjects. Both interesting research venues are left open for future contributions.

¹The range of apparent scene depths as seen by the viewer in a projected stereoscopic image/video.

1.1 Background of the idea

One of the simplest approaches to reduce the eye strain is to use narrow baselines in acquisition time. A gentler stereo produced by shooting with a narrow baseline or small 3D budgets has the effect of limiting the display to a safe parallax range for visualization [4, 13]. The obvious downfall of this approach is the lack of apparent *depth richness* that may not appeal to some viewers.

Non-linear image warping has been used to adapt visual content to different displays [14] and a variation of this idea has been proposed to eliminate disparities outside the comfort tolerance [15]. This method is based on a warping transformation defined by a non-linear transfer function applied to the disparity map. The novel stereoscopic image created by this method is a depth distorted version of the original that is visually comfortable [16].

Limiting the image DOF by optical methods by using large camera apertures, or by out-of-focus emulation in post-processing has been used to enhance the viewing comfort. This idea is that our binocular vision system attempts to *stereo-fuse* (to extract depth from parallax views of a scene) as long the image is sharply focused [17, 18, 5]. This mechanism prevents eye strain by ignoring scenic elements outside the viewed region-of-interest (ROI) by means of adjusting the eye DOF (accommodation) in order to render sharply only the part of the scene we are interested in seeing; in other words, *out-of-focus blurring* can be used to prevent visual discomfort caused by regions with excessive disparities in the projected scene [19, 20, 21, 22].

Methods that relies on vergence eye movement emulation to prevent eye strain has been proposed for immersive 3D media applications such as virtual reality [23]. The screen parallax adjustment based on controlling the disparity of feature points has been proposed by Chen et. al in [24] as visual discomfort palliative. The effect of affecting the image disparity as a mean to emulate vergence and control the screen parallax over the ROI has been presented in [25] by Bernhard et al. who studied the time lapses necessary to produce stereofusion.

Some of these proposals treat the vergence and DOF as separated problems; in this paper, I propose to consider simultaneously both solutions for the visual discomfort.

1.2 Proposed method

The method presented in this paper aims to emulates the vergence eye movement by exploiting the viewer's gaze direction feedback provided by an eye-tracking device. Furthermore, the method emulates the accommodation of the eyes by controlling the DOF of the stereoscopic image. This two-fold approach is implemented after the stereo imagery has been acquired and relies on a depth-based segmentation of the scene. A dense disparity map of each stereoscopic image must be calculated in post-processing or obtained alongside the stereoscopic imagery by means of time-of-flight cameras or range sensors.

Once the viewer's gaze direction over the screen is tracked, a disparity correction is done by means of a simple image horizontal alignment transformation. This transformation has the effect of bringing the ROI in the stereoscopic image to the screen distance (zero parallax). Since eyes are focused on the screen distance, this method naturally eliminates the vergence-accommodation conflicts than hinder displaying 3D imagery on planar screens. After the correction, the peripheral of ROI will contain enough binocular information to perceive depth, but without visual strain.

The vergence correction may produce excessive disparities regions outside the ROI, which are corrected by emulating out-of-focus blurring in selected regions. This spatially dependent filtering limits the scene's DOF, preventing the stereo-fusion of areas with excessive screen parallax. Both corrections, vergence and accommodations, should be implemented dynamically, every time the viewer's gaze direction changes. Special attention should be taken to the potentially adverse effect

in the visual comfort of fast viewing direction transitions. This topic is further discussed in Section 5.

This technique has the advantage of producing a more natural stereoscopic viewing on planar screens without introducing depth distortions as other methods based on image warping. In addition, this idea has a clear application to personalized visualization of image-based stereoscopic media where 3D videos can be streamed to personal visualization devices such as head-mounted displays with eye tracking capabilities.

The next section of this paper introduces the vergence-accommodation synergy and how it affects visual comfort in the context of stereoscopic content projected on planar screens. The proposed image corrections are explained in Section 3 followed by examples in Section 4 to illustrate the benefits of this technique.

2 Visual comfort

Stereopsis is a perceptual phenomena that occurs when two different views of the same scene induce a subjective perception of depth [26]. In real-world viewing scenarios, there is no discrepancy between where the eyes converge in the scene and their distance of focusing; both binocular mechanisms work cooperatively to experience the world's depth. However, stereoscopic imaging displayed on planar screens may introduce discrepancies between focus and vergence distances. These conflictive cues are one source of eye fatigue, specially when the display is close to the viewer.

2.1 Vergence

The vergence eye movement assures that the same ROI in the scene is imaged on the Panum's area of the retina in both eyes [27]. An example that illustrates the binocular vergence is in Fig. 1, where eyes verge at infinity (example on the left), so objects in the ROI exhibit maximum screen parallax. This situation can be modeled as two pin-hole cameras separated by a baseline and with parallel optical axes. In the right side example, eyes have to toe in (vergence) to overlap the statue (the ROI) on the same retinal region in both eyes. Vergence makes scene outside the ROI double, which leads to eye strain; but eye focus (accommodation) changes in order focus the ROI.

2.2 Accommodation

Accommodation is the mechanism by which eyes focus the image projected on each Panum's area in order to obtain maximum visual acuity on the ROI. This mechanism is closely tied to the vergence eye movement. For instance, when our eyes verge to a point in the scene, some elements in the scene may show excessive horizontal disparities.

The natural accommodation of the eyes defines the appropriate DOF to visualize the objects at a given distance, while scene elements outside that range appear doubled and blurred. Our mind processes only the sharply focused part of the scene by using the slight image differences that arise from parallax views. This mechanism is called *stereo-fusion* and is in part responsible of our binocular perception of depth. Perceiving parts of the scene blurred prevents stereo-fusion of scenic elements that exhibit horizontal disparities outside the comfortable range [10].

The accommodation of the eyes to different depths can be illustrated using a monoscopic picture acquired with narrow DOF, e.g., the one produced by a large camera aperture (small f-stop). In the example shown in Fig. 2, the focus is on scenic elements at different distances using a single camera to illustrate the effect of a narrow DOF; but this example does not show the double image effect seen by converging our eyes towards the same ROI as shown in Fig. 1.

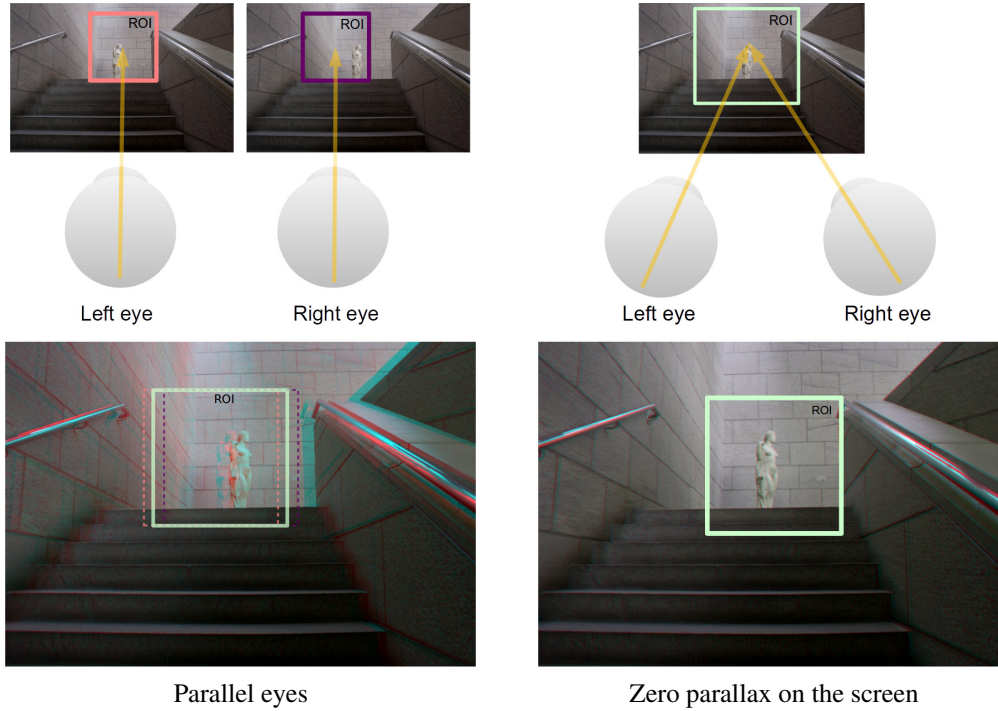


Figure 1: Two different projections of a scene captured with the same stereoscopic camera arrangement. In the left side example, vergence is at infinity (parallel optical axis), while in the case depicted in the example on the right, vergence is on the target object in the ROI. Stereoscopic views are shown at the bottom of each example as red-cyan anaglyphs.

The mechanisms illustrated in Fig. 1 and 2 happen automatically during our daily routine and make possible the binocular part of our depth perception. The challenge is to incorporate these mechanisms in situations where binocular vision is emulated by watching a planar screen instead to a volumetric area.

2.3 Vergence-accommodation discrepancy

The tolerance range between focusing and verging distances for a comfortable stereoscopic visualization has been previously studied [9, 5, 4, 28]. A model used to derivate the tolerable screen parallax boundaries for a planar display at certain distance of the viewer is shown in Fig. 3. This model represents two eye with interocular baseline b that watches two projections of the same virtual object on a screen located at distance a . The virtual object exhibits a screen parallax p as seen by each eye. This screen parallax forces the verge eye movement determined by the distance v , but eyes must focus on the screen, which is located at the accommodation distance a . The difference between vergence and accommodation distances is derived by

$$p/(b \cdot a) = (1/a - 1/v). \quad (1)$$

Note that, for a baseline b , the same stereoscopic content will stimulate similar depth in different displays as long the ratio p/a is constant. For instance, viewers rank high in visual comfort when



Figure 2: Eye accommodation at different distances from the observer. In a sharply rendered stereoscopic scene, the viewer’s eyes naturally accommodate to focus the screen or display.

3D movies are properly re-targeted for television displays [3], despite the eye-screen distance in the living room is much smaller than in a movie theater. For a given b , the ratio of distances p/a can be excessively large, inducing eye strain. A typical example of this is found when 3D imagery acquired with large stereoscopic baselines is displayed at small screen distances, such as in head-mounted displays or computer monitors.

The tested accommodation-vergence discrepancies range to avoid visual discomfort was presented by Lambooj et al. [9] and tested by Shibata et al. [26] among other authors.

3 Method

Vergence can be emulated by tracking the viewer’s gaze direction and by using this information to define the ROI’s coordinates in the stereoscopic image. This feedback information is used to dynamically align the image pair to produce a region of zero horizontal disparity over the ROI, which would appear as a zero parallax area on the screen. This scheme simulates the natural process of vergence, eliminating any vergence-accommodation conflict since the ROI is always rendered at screen distance where the eyes are focused. An example to illustrate this concept is illustrated in Fig. 1.

This approach may introduce excessive horizontal disparities in areas outside the ROI region. This is where a spatially selective blurring is necessary in order to prevent stereo-fusion of image regions that would otherwise produce eye strain [28, 10].

3.1 Disparity maps

Each image must be segmented based on the disparity exhibited by each scene element after the vergence correction. This is necessary in order to identify which image regions are within the desired DOF and which areas need to appear out-of-focus as part of the accommodation correction.

An stereoscopic image pair (I_L, I_R) defines a pair of dense disparity maps (D_L, D_R) . The Euclidean distance between corresponding pixels in I_L and I_R is

$$D_L(u, v) = d_L, \quad (2)$$

which defines the correspondence

$$I_L(u, v) = I_R(u - d_L, v). \quad (3)$$

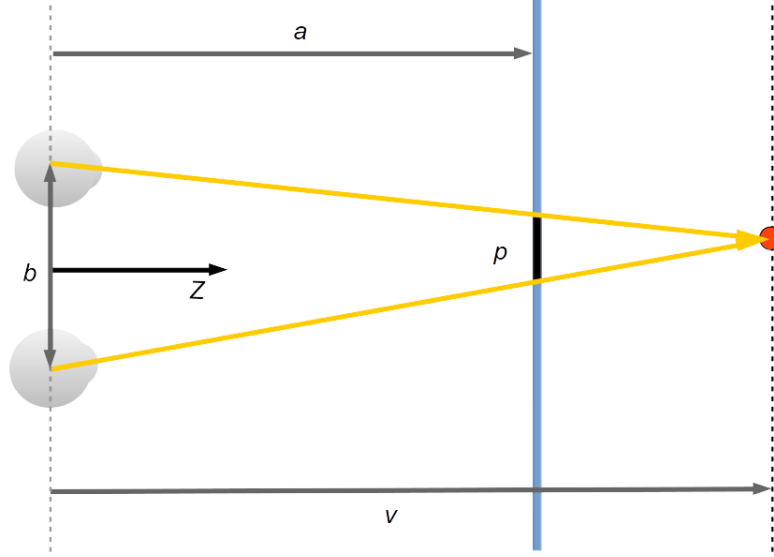


Figure 3: Geometric relationships involved in the stereoscopic view of a virtual object projected on a planar screen. A screen parallax p forces viewers with an interocular distance b to toe in their eyes (vergence eye movement) in such an angle that makes the virtual object located a distance v be imaged simultaneously on the Panum's area of each eye. At the same time, the viewers' eyes are forced to focus on a screen located at distance a , e.g, accommodating the eyes so the screen appears in the visual DOF.

Conversely, $D_R(u', v')$ defines the Euclidean distance d_R , which represents the correspondence

$$I_R(u', v') = I_L(u - d_R, v'). \quad (4)$$

Overlooking for a moment the effect of occlusions and matching errors, D_L contains a one-to-one pixel matching from $I_L \rightarrow I_R$ and D_R contains a pixel matching from $I_R \rightarrow I_L$. In this scenario,

$$d_L = -d_R, \quad (5)$$

for any pair corresponding points in (I_L, I_R) . In this context, D_L and D_R will be used to segment respectively I_L and I_R into distinctive depth layers.

3.2 Gaze direction

The user's gaze direction defines a point with coordinates (U, V) in the planar screen. This point can be retro-projected on each image plane as a point with coordinates (u, v) and be used to define which scenic element the user is watching. This can be done by choosing one of the images in (I_L, I_R) as a reference for gaze tracking, e.g, if the left image in the set is chosen as the tracking image, (U, V) is retro-projected onto a point with coordinates (u_L, v_L) in image I_L .

3.3 Vergence correction

Continuing with the same example, the coordinates (u_L, v_L) obtained from the viewer's gaze direction are then used to determine the Euclidean distance between the point $I_L(u_L, v_L)$ and the corresponding point in I_R . This Euclidean distance d , defined by the disparity map $D_L(u_L, v_L)$, is the

amount to translate I_R to emulate the eye vergence on a ROI defined by (u_L, v_L) . Formally,

$$I_{Rc} = I_R(u - d, v). \quad (6)$$

This transformation guarantee that the set (I_L, I_{Rc}) exhibits zero disparity on the ROI centered in (u_L, v_L) . The alignment between I_L and I_R is maintained by applying a translation $I_R(u - d, v)$ every time d is updated, which happens every time the gaze direction changes. Although the center of ROI shows zero screen parallax after correction, its peripheral will maintain a richness of horizontal disparities enough to stimulate binocular depth.

3.4 Disparity map correction

The distance d is used to correct the disparity maps of vergence, which has to be done before the actual image alignment. This is done to segment (I_L, I_R) into two mutually exclusive regions: one represent the on-focus scene, which is within a defined DOF for comfortable viewing, and a region outside this DOF, which needs to be processed. The corrected disparity maps are

$$\begin{aligned} D_{Lc} &= D_L(u - d, v), \text{ and} \\ D_{Rc} &= D_R(u + d, v), \end{aligned} \quad (7)$$

where d is defined by $D_L(u_l, v_l)$ when I_L is used as reference image to track the viewer's gaze direction.

3.5 Disparity map generation

The disparity maps can be obtained by calculating the optical flow between pair of stereoscopic images, by manually segmenting each image based on a depth model of the scene, or by acquiring dense disparity maps by means of depth sensors.

A post-acquisition approach is to calculate D_L and D_R from the optical flow between I_L and I_R . A coarse image segmentation based on depth can be implemented by using this method. But, edges in most targets may not be well defined. Occlusions and matching errors make this option inviable in most cases; at least, disparity maps would not be usable without a previous data correction. A fully automatic image segmentation based on depth is a problem open for further research.

Alternatively, a semi-automatic method to define a depth-based segmentation on both images can be considered. This approach consists in selecting N visual targets in I_L and I_R , manually or automatically segment these targets and calculate their disparities d_i ($i \in \{1, \dots, N\}$). Then, define a model for the scene depth and create a corresponding disparity background. Finally, combine target and background disparities to synthesize D_L and D_R .

Perhaps the best approach is to acquire the stereoscopic imagery alongside depth information of the scene by means of time-of-flight cameras or range sensors. Using this approach, which is technologically available today, would greatly improve the necessary information to implement the proposed corrections every time the displaying scenario render them necessary.

3.6 Simulated accommodation

The out-of-focus emulation in selected image regions is implemented by applying a spatially selective low-pass filter to each image. Blohm et al. proposed an out-of-focus blurring model suitable for this problem [12]. Their approach is based on the thin-lens model where rays are de-focused based on the target's relative depth from an on-focus optimal distance z_o . In this problem, the optimally

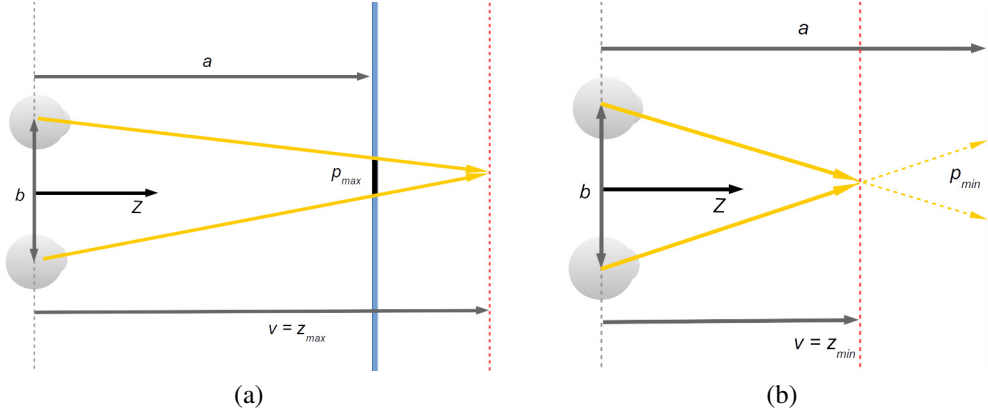


Figure 4: The maximum distance z_{max} , determined by the maximum eye vergence, produces a positive screen parallax p_{max} , as shown in (a), which makes the scene appear *behind* the screen (positive space). Conversely, a minimum vergence distance z_{min} produces a negative parallax in the screen p_{min} , as shown in (b), which conveys the illusion of being *in front* of the screen (negative space).

focused distance is a (the screen); hence, objects behind ($z > a$) and in front ($z < a$) of the screen will be blurred based on their z distance. Therefore, the selective image blurring will depend on the disparity after vergence correction.

The blurring is done by a low-pass two-dimensional filter defined by

$$h(u, v, r) = \begin{cases} \frac{1}{\pi \cdot r^2} & \text{if } \sqrt{u^2 + v^2} \leq r \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where r is the *circle of confusion* radius and is defined as

$$r = \frac{p_d \cdot a}{2} \cdot \left(\frac{1}{z} - \frac{1}{a} \right), \quad (9)$$

in this equation, p_d is the pupil's aperture diameter and z is the distance to virtual scene with respect to the viewer (Fig. 3), which is defined as a function of the screen parallax p , is given by

$$z = \frac{a \cdot b}{b - p}. \quad (10)$$

The last equation can be used to obtain the parallax limits $[p_{min}, p_{max}]$, which are defined by the DOF for comfortable viewing. The geometric relationship that originates the parallax range are illustrated in 4. The screen parallax range must be converted to a horizontal disparity range $[d_{min}, d_{max}]$ by using a projective transformation in order to find the DOF regions in from (D_{Lc}, D_{Rc}) .

3.7 Screen distances and DOF

The pupil's diameter p_d affects the DOF according to the thin lens model. This parameter can varies between subjects depending of their age and gender; for instance, the average p_d for the adult population varies in the 2 – 10 mm range [29]. In our test, the pupil diameter was fixed to 10 mm to match the light conditions in the displaying room and the chosen DOF was 0.2 diopters [3].

These settings produces a comfortable stereoscopic visualization in the tested display; however, these settings must be adjusted for a different displaying scenario.

The accommodative eye DOF is adaptive during normal viewing; but, in this context, its value is fixed since the screen distance is also fixed. In this example, the chosen DOF determines a strain-free visualization between

$$\begin{aligned} z_{min} &= 0.9 \cdot a, & \text{and} \\ z_{max} &= 1.1 \cdot a. \end{aligned} \tag{11}$$

3.8 Summary: stereoscopic image correction

The correction process can be applied to any stereoscopic pair of images (I_L, I_R) acquired with a panoramic DOF that guarantee focal acuity over the whole scene. Acquisition parameters such as the stereoscopic baseline does not to be known for the correction; only displaying parameters such as screen distance, projected image aspect ratio and size are required. The pupil diameter p_d and desired DOF must be tuned to the displaying conditions.

First, a target ROI is defined by calculating the image coordinated (u, v) obtained by tracking the viewer’s gazing coordinates in the screen (U, V) . Any of the images in (I_L, I_R) can be used as a reference; I_L is used in this example.

The disparity maps D_L and D_R are obtained by any of the alternatives explained in Section 3.5.

The horizontal disparity d at the ROI’s center, which has coordinates (u, v) , is obtained from $D_L(u, v)$.

The vergence correction is applied to both disparity maps as defined in Eq. 7 generating D_{Lc} and D_{Rc} . These corrected disparity maps are used to define two regions in each image: one is within the DOF and will appear on-focus and the other is subject to out-of-focus blurring.

Both images are then subject to the spatially selective low-pass filtering defined in Eq. 8, 9 and 10, to create the accommodation corrected version of the source image pair:

$$\begin{aligned} I_{La} &= I_L * h, & \text{and} \\ I_{Ra} &= I_R * h, \end{aligned} \tag{12}$$

where $*$ represents the convolution between an image and the and the spatially selective, low pass-filter h .

Finally, the vergence alignment is applied to I_R as a translation transformation:

$$\begin{aligned} I_{Lc} &= I_{La}, & \text{and} \\ I_{Rc} &= I_{Ra}(u - d, v). \end{aligned} \tag{13}$$

Similar procedure can be deducted by using I_R instead of I_L as the reference image for tracking the viewer’s gaze direction.

4 Experiments

The correction method was tested in a series of hand-colored stereoscopic prints, a.k.a. *stereo-views*, from circa 1905 [30], which, in addition to their inherent historical value, have interesting characteristics for this test. These scenes were not intended for virtual scene navigation, but to be visualized as a single scene with the assistant of a special purpose viewer, which position each image a few centimeters from the viewer’s eyes². Visual strain is evident when these images are displayed

²Alternatively, binocular depth illusion can also be achieved by watching these stereo-views using the cross-eye method, which requires locating the stereo-view a few centimeters from the eyes.

at a close distance from the viewers’ eyes and the gaze is directed towards close or far regions of the virtual scene. This problems arises in some stereo-views because they were acquired with slightly too large baselines for the intended screen distance.

4.1 Test parameters

As part of the experimental setup, the chosen baseline b was 70 mm, which corresponds approximately to the average interocular separation in the adult human population. The display distance a was 50 cm and the image height was 26 cm in all the tested cases.

The DOF was set to 0.2 dioptres, which determines the apparent minimum and maximum depth of objects rendered sharply on screen; in this case, the distance z_{min} is 475 mm and z_{max} is 550 mm. The depth tolerance to maintain a comfortable binocular visualization is tied to this acceptable screen parallax limits, which were derived from z_{min} and z_{max} as p_{min} (−22.8 mm) and p_{max} (18.7 mm), respectively. These limits are shown in Fig. 4.

In this test, a pupil aperture diameter p_d (10 mm) was found experimentally to be appropriate for emulating the natural viewing conditions in the experimental setting. A smaller p_d would determine a larger DOF; hence, most of the scene will appear sharply rendered, which may lead to visual discomfort for the screen distance of these examples. Conversely, a larger p_d would lead to a narrower DOF, which would introduce excessive blurring in some images areas; the latter produce an unnatural effect.

Note that by maintaining the same DOF and moving back the display to a larger a , or by maintaining the display distance but reducing the stereoscopic image size, the whole image could be visually explored by only correcting vergence since the whole stereoscopic projection will be between the acceptable DOF.

4.2 Disparity map

Given the quality of source images, the optical flow method to obtain (D_L, D_R) was not practical. Therefore, the disparity maps for this test were manually generated (Section 3.5) by selecting the N most likely targets in each scene. An example using I_L is shown in Fig. 5.

4.3 Viewer’s gaze feedback and disparity-map correction

It was not necessary to count on eye tracking device to test the correction method; it was enough to emulate the gaze direction by manually selecting a few candidate targets and extracting their image coordinates. In this fashion, two extreme gazing directions were selected to apply the correction technique: one emulates looking at the closest element in each scene and the other emulates gazing at the farther scenic elements. For instance, in the example illustrated in Fig. 6-(a), the viewer’s gaze at the closer element in I_L , which is indicated as a frame whose center is located in (u_1, v_1) . The horizontal disparity at this gaze direction was calculated from $D_L(u_1, v_1)$, which is labeled d_1 .

Before aligning the stereoscopic images, d_1 was used to correct the disparity maps as indicated in Section 3.4 to generate D_{Lc} and D_{Rc} . The disparity limits $[d_{min}, d_{max}]$ were calculated based on Eq. 10 and 1.

Finally, masks for I_L and I_R were created based on the disparity limits to indicate which image regions were within the DOF and which needed to be blurred. These masks, (M_L, M_R) , were assigned the value 0 whenever $D_{jc}(u, v) \in [d_{min}, d_{max}]$, ($j \in \{L, R\}$), and have value 1 otherwise. In Fig. 6-(b), a gray scale representation of M_L is shown, indicating in color black the regions in I_L that needed out-of-focus blurring.

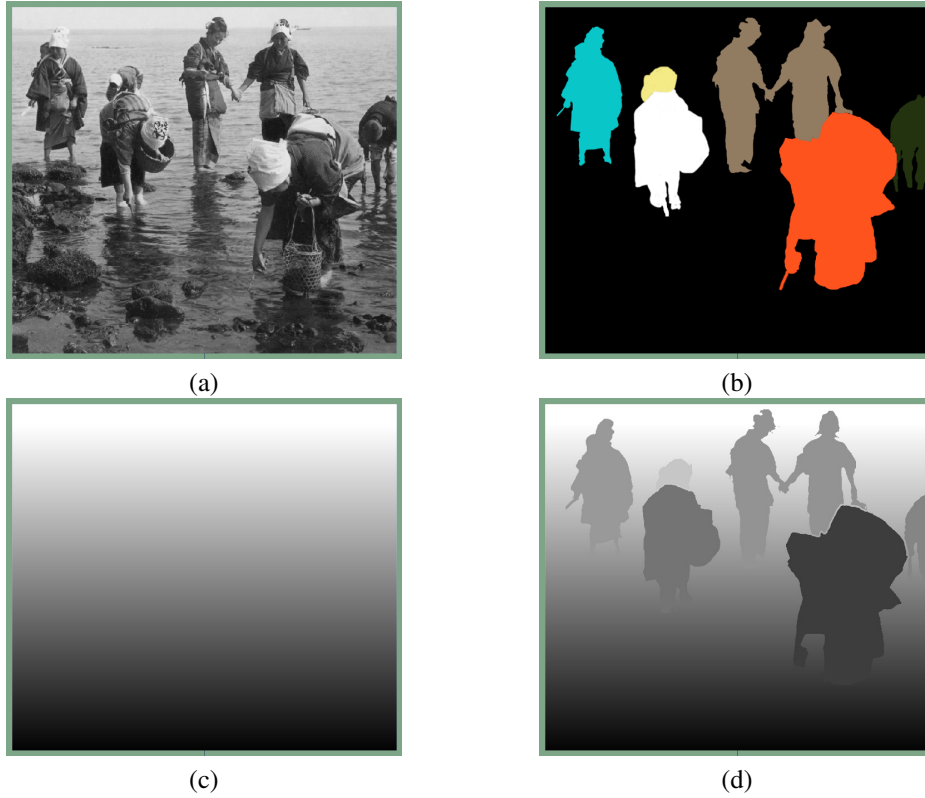


Figure 5: Generation of the disparity map of the scene. The example shows I_L (a) and the segmentation of the N most likely targets (b). The background disparity map D_b (c) was generated based on a disparity model for this particular scene. Finally, the generated D_L (d) was created by using D_b and the target disparities d_i . The gray scale colors encodes increasing disparities starting from a minimum horizontal disparity at the bottom (darkest color) to its maximum value at the top.



Figure 6: Gazing direction and *on-focus* vs. *out-of-focus* regions in I_L . The ROI (a) is centered in (u_1, v_1) and defines the amount of horizontal translation correction d_1 , which was calculated from $D_{Lc}(u_1, v_1)$. The blurring mask (b) showing *out-of-focus* regions (black) and *on-focus* (white) for the chosen ROI.



Figure 7: The out-of-focus emulation. Result of spatially selective filtering I_{La} (a), and details of the boundaries between region with large disparity differences (b).

4.4 Accommodation correction

Once defined which regions should be blurred, the spatially selective filter (Eq. 8) was applied to each pixel in (I_L, I_R) . The amount of defocussing (Eq. 9) was based on each the region or pixel disparity, which was obtained from $D_{Lc}(u, v)$ and $D_{Rc}(u, v)$. The accommodation corrected image I_{La} is presented in Fig. 7-(a).

The edge of each segmented target was smoothed by applying a Gaussian low-pass filter in order to improve the blending between blurred regions with large disparity gradients. A detail of this effect is shown in Fig. 7-(b).

4.5 Results

The result of using the corrective method in two different stereo-views are presented in Fig. 8 and 9. These results are shown as red-cyan anaglyphs and were produced to be displayed in the particular settings defined for this test, i.e., a display distance of 50 cm and image height of 26 cm. It would be advisable to zoom into each result if possible in order to emulate the intended displaying conditions.

The left columns of images show the original stereoscopic sets, where the target ROI is indicated with a rectangular frame. Right columns contain the final vergence-accommodation (I_{Lc}, I_{Rc}) for each ROI. In both cases, the upper rows of images correspond to a ROI in the closer region of the scene, whereas in the bottom rows of images, the ROI is in the farthest region of the scene.

5 Discussion

A potential source of visual discomfort may appear as result of fast depth transitions while navigating the scene. More specifically, the introduced vergence correction may produce discomfort by perceiving the whole scene moving back and forth with respect to the screen. However, the DOF correction will limit stereofusion to regions of the scene with admissible disparities and could tamper this potential issue. A good strategy to deal with this problem would be to smooth the speed at which the vergence alignment and DOF correction are applied [25]. This aspect of the corrective algorithm needs to be studied in detail with subjective trials, which is an interesting venue for future research.

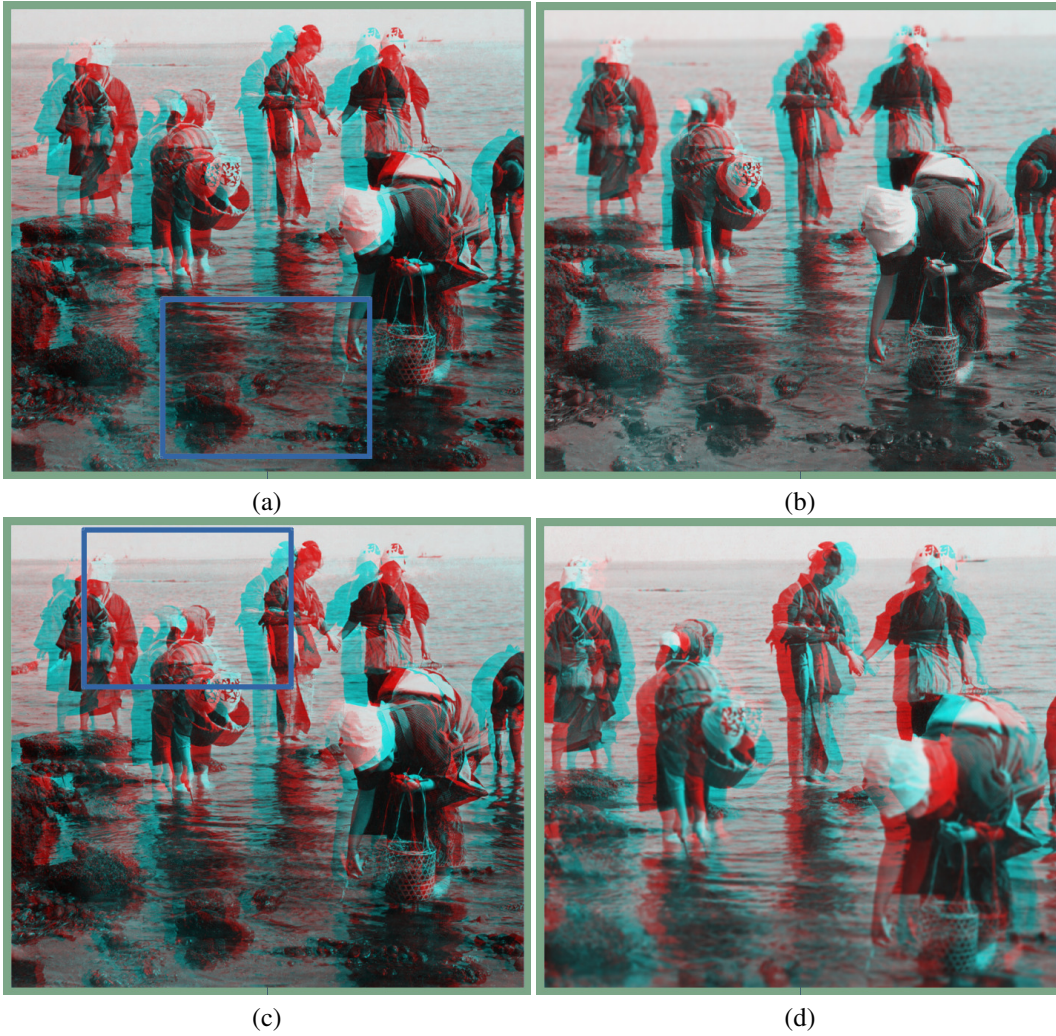


Figure 8: Example of vergence-accommodation correction [31]. In (a), the ROI is in the original stereoscopic image bottom (closest point of the virtual scene to the viewer), whereas in (c), the ROI is located in the original image's top (the horizon in this scene). The visual comfort improvement for each ROI is shown in (b) and (c).

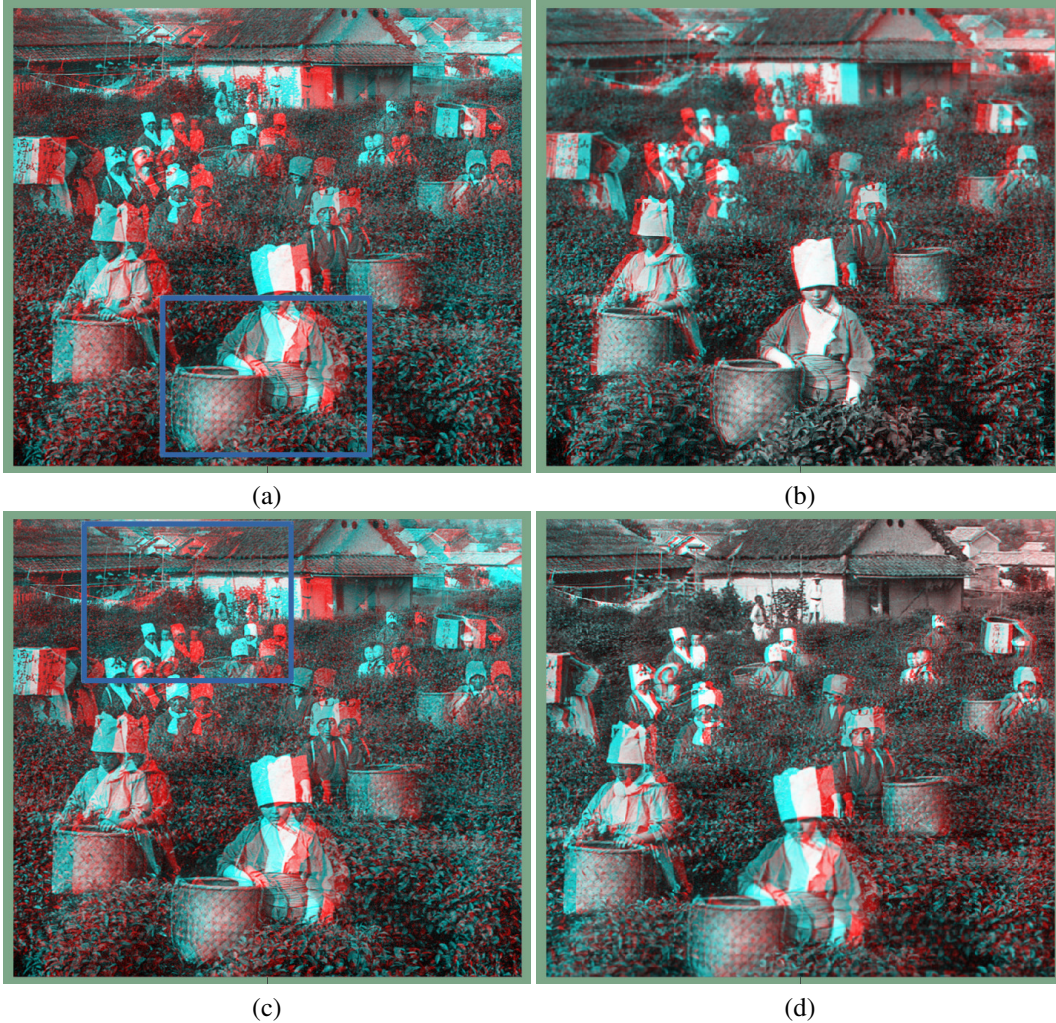


Figure 9: Another example the vergence-accommodation correction [32]. In (a), the ROI is in the original image bottom, e.g., the closest tea's picker to the viewer in this virtual scene, whereas in (c), the ROI is located in the original image's top, e.g., houses in the horizon line. The method results are illustrated in (b) and (d).

6 Conclusions

This paper presents a technique to correct stereoscopic imagery to be displayed in close eye-screen situations. The correction method emulates binocular mechanisms of vergence and accommodation to prevent visual discomfort and relays on the feedback of the viewer's gaze direction. Unlike other approaches that distort images in the stereoscopic set to eliminate excessive horizontal disparities, the approach presented herein does not alter the depth relationship in the projected scene; instead, it emulates binocular vision mechanisms to counteract some inherent limitations of displaying stereoscopic imagery on planar screens. One important aspect that needs further research is the depth-based segmentation of each image. In order to illustrate the corrective method, manually generated disparity maps were used in this paper. A fully automatic approach to perform an depth-based image segmentation can rely on the combined use of color information and optical flow calculations, or in the input of depth-sensors. This does not hinder the potential use of this method to produce a more comfortable visualization of immersive stereoscopic environments based on real-world imagery.

References

- [1] Cassin B, Rubin ML, Solomon S (1984), *Dictionary of Eye Terminology*. Triad Publishing Company
- [2] Lambooij M, IJsselsteijn W, Bouwhuis, DG, and Heynderickx, I (2011) Evaluation of stereoscopic images: beyond 2D quality. *IEEE Tran. on Broadcasting*, 57(2):432–444
- [3] Read JC (2014) Viewer experience with stereoscopic 3D television in the home. *Displays* 35(5):252–260
- [4] Siegel M, Tobinaga Y, Akiya T, Merritt JO, Bolas MT, and Fisher SS (1999) Kinder, gentler stereo. *Proc. of SPIE* 3639:18–27
- [5] Wöpking M (1995) Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus. *J. Soc. Info. Display* 3(3):101–103
- [6] Siegel M, Nagata S (2000) Just enough reality: comfortable 3-D viewing via microstereopsis. *IEEE Tran. on Circuits and Systems for Video Technology* 10:387–396
- [7] Gurrieri LE, Dubois E (2014) Depth consistency and vertical disparities in stereoscopic panoramas. *J. of Elec. Imaging* 23(1):011004. DOI 10.1117/1.JEI.23.1.011004
- [8] Bando T, Iijima A, and Yano, S (2012) Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: A review. *Displays* 33(2):76–83
- [9] Lambooij, M, IJsselsteijn W, Fortuin M., and Heynderickx I (2009) Visual discomfort in stereoscopic displays: a review. *J. of Imag. Sci. and Tech.* 53(3):030201–14
- [10] Howarth PA (2011) Potential hazards of viewing 3-D stereoscopic television, cinema and computer games: a review. *Ophthalmic and Physiological Optics* 31(2):111–122
- [11] Hillaire S, Lécuyer A, Cozot R, Casiez G (2008) Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments. *IEEE Virtual Reality Conf. VR'08* 47–50.

- [12] Blohm W, Beldie I, Schenke K, Fazel K, Pastoor S (1997) Stereoscopic image representation with synthetic depth of field. *J. Soc. Info. Display* 5(3):307–313
- [13] Gurrieri LE, Dubois E (2013) Stereoscopic cameras for the real-time acquisition of panoramic 3D images and videos. *Proc. SPIE Stereo.scopic Displays and App. XXIV 86481W*:1–17. DOI 10.1117/12.2002129
- [14] Shamir A, Sorkine O (2009) Visual media retargeting. *ACM SIGGRAPH ASIA 2009 Courses* 1:1–11
- [15] Lang M, Hornung A, Wang O, Poulakos A, Smolic A, Gross M (2010) Nonlinear disparity mapping for stereoscopic 3D. *ACM Tran. on Graphics* 29(4):75–85
- [16] Wang O, Lang M, Stefanoski N, Sorkine-Hornung A, Sorkine-Hornung O, Smolic A, Gross M (2013) Image domain warping for stereoscopic 3D applications. *Emerging Technologies for 3D Video* 207–230
- [17] Subedar MM, Karam LJ (2010) A study of relation between blur and depth in stereoscopic images. *Int. Workshop Video Process Quality Metrics for Consumer Electronics* 1–4
- [18] Sun G, Holliman N (2009) Evaluating methods for controlling depth perception in stereoscopic cinematography. *Proc. of SPIE Int. Soc. for Optics and Photonics* 7237:72370I
- [19] Dias JM, Araujo H, Batista JE, de Almeida AT (1992) Depth perception by controlling focus. *Sensor Fusion IV: Control Paradigms and Data Structures* 1611:214–224
- [20] Leroy L, Fuchs P, Moreau G (2010) Real-time comfort enhancement in stereoscopic displays by disparity and content-adapted blur *IEEE Int Symp. on Industrial Electronics* 3595–3600
- [21] Hillaire S, Lécuyer A, Cozot R, Casiez G (2008) Depth-of-field blur effects for first-person navigation in virtual environments. *IEEE Computer Graphics and App.* 28(6):47–55.
- [22] Blum T, Wieczorek M, Aichert A, Tibrewal R, Navab N (2010) The effect of out-of-focus blur on visual discomfort when using stereo displays *IEEE Int. Symp. on Mixed and Augmented Reality* 13–17
- [23] Sherstyuk A, Dey A, Sandor C, et. al. (2012) Dynamic eye convergence for head-mounted displays improves user performance in virtual environments. *Proc. of the ACM SIGGRAPH Symp. on Interactive 3D Graphics and Games* 1:23–30
- [24] Chen WC, Hsiao FJ, Lin CW (2010) An automatic parallax adjustment method for stereoscopic augmented reality systems. *IEEE Int. Symp. on Mixed and Augmented Reality* 215–216
- [25] Bernhard M, Dell'mour C, Hecher M, Stavrakis E, Wimmer M (2014) The effects of fast disparity adjustment in gaze-controlled stereoscopic applications. *Proc. of the Symp. on Eye Tracking Research and App.* 111–118
- [26] Shibata T, Kim J, Hoffman DM, Banks MS (2011) The zone of comfort: Predicting visual discomfort with stereo displays. *J. of Vision* 11(8):1–11
- [27] Steinman S, Steinman B, Garzia R (2000) *Foundations of Binocular Vision: A Clinical Perspective*. McGraw-Hill Medical, New York, NY

- [28] Hoffman DM, Girshick, AR, Akeley, K, Banks MS (2008) Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *J. of Vision* 8(3):1–30
- [29] Winn B, Whitaker D, Elliott DB, Phillips NJ (1994) Factors affecting light-adapted pupil size in normal human subjects. *Investigative Ophthalmology & Visual Science* 35(3):1132–1137
- [30] Okinawa Soba (Rob) (2015) T.Enami’s stereoviews of Japan available on-line under Creative Commons: Attribution - Non-Commercial - Share Alike, 2.0. Publishing Flickr <https://flic.kr/s/aHsj9paATA>. Accessed 3 Jun 2015
- [31] T. Enami (ca.1905) Shell Pickers at Honmoku Near Yokohama available on-line under Creative Commons: Attribution - Non-Commercial - Share Alike, 2.0. Publishing Flickr <https://flic.kr/p/4wifes>. Accessed 19 Jun 2015
- [32] Herbert G. Ponting (ca. 1904) SOME NICE STEREOVIEWS OF OLD MEIJI-ERA JAPAN - From a 100-View Set by Underwood & Underwood available on-line under Creative Commons: Attribution - Non-Commercial - Share Alike, 2.0. Publishing Flickr <https://flic.kr/p/4F9NBi>. Accessed 19 Jun 2015