Steady-State Feedback Analysis of Tele-Graffiti

Naoya Takao
Matsushita Research
Matsushita Corporation
Osaka, Japan

Simon Baker
The Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Jianbo Shi
The GRASP Laboratory
University of Pensylvania
Philadelphia, PA 19104

We analyze the feedback loop in Tele-Graffiti, a camera-projector based remote sketching system which we recently developed. We derive the gain through the feedback loop and the final images that will be viewed by the users of the system. We then derive the "optimal gain" as the gain that results in the final viewed images being as close as possible to the sum of the images actually drawn on the paper. We also propose a way of using the projector to compensate for insufficient ambient light and describe how our analysis is affected by: (1) automatic gain control (AGC) and (2) color imagery. We end by proposing an algorithm to decompose the final viewed images into estimates of the drawn images.

1 Introduction

Systems combining video cameras and LCD projectors are becoming more and more prevalent. The first such system was Pierre Wellner's Xerox "DigitalDesk" [Wellner, 1993]. Other such systems include the University of North Carolina's "Office of the Future" [Raskar *et al.*, 1998], INRIA Grenoble's "MagicBoard" [Hall *et al.*, 1999], and Yoichi Sato's "Augmented Desk" [Sato *et al.*, 2000]. Recently, there has been a growing body of work in the computer vision literature aimed at using video cameras to improve the use of projectors as display devices [Raskar and Beardsley, 2001, Sukthankar *et al.*, 2001a, Sukthankar *et al.*, 2001b].

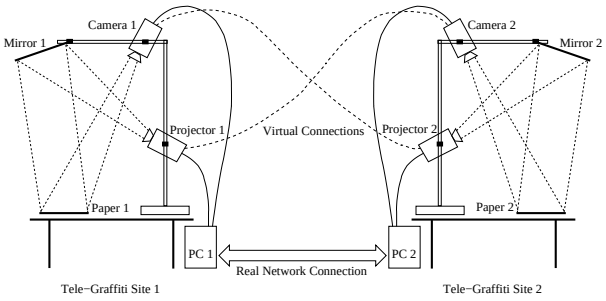
Recently we have developed "Tele-Graffiti," a camera-projector based remote sketching system [Takao *et al.*, 2003]. Tele-Graffiti allows two people to communicate remotely via hand-drawn sketches. What one person draws at one site is captured using a video camera, transmitted to the other site, and displayed there using an LCD projector. See Figure 1(a) for a schematic diagram of two Tele-Graffiti sites connect by a network and Figure 1(b) for an image of a real Tele-Graffiti system. See the accompanying movie "telegraffiti-uidemo.mpg" for an example of two people interacting using Tele-Graffiti. Also, see [Takao *et al.*, 2002] for larger versions of the figures in this paper.

The combination of projectors and cameras in the same system can either be beneficial or a hindrance, depending on what you want to do. In hand-tracking systems for desktop user-interfaces, the light projected by the projector can make simple background subtraction algorithms inapplicable. Instead, several groups have used infra-red cameras for hand-tracking, a modality unaffected by the light radiated by the projector [Sato *et al.*, 2000, Leibe *et al.*, 2000]. On the other hand, the projector can be used to project structured light, thereby allowing the estimation of 3D shape using the "laser range finding" principal. By carefully synchronizing the camera and the projector, the structured light can even be time-multiplexed with the projected image to make it imperceptible to humans [Raskar *et al.*, 1998].

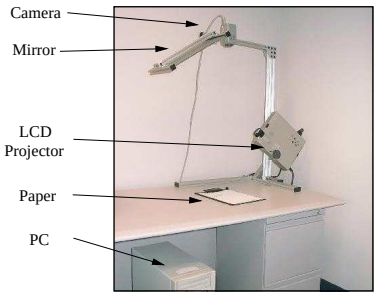
A particular problem in certain applications is *feedback*. If the projector radiates light that is then imaged by the camera, and the resulting image then passed back to the projector, potentially after being processed in one way or another, there is the potential for feedback. It is possible that the projected image may get brighter and brighter until the camera saturates resulting in an unusable system. Alternatively, the system might oscillate between two states. Finally, there is the possibility of "visual echoing" [Takao *et al.*, 2002]. On the other hand, feedback can be used to help the system, as is done in [Sukthankar *et al.*, 2001a] for shadow removal.

Camera-projector based communication systems are all susceptible to feedback. When we first gave a demonstration of Tele-Graffiti at ICCV [Takao *et al.*, 2001], the system initially suffered from severe feedback, quickly saturated, and became unusable. At the time we added extra lights as a "quick fix". In this paper we study the cause of the feedback and propose a principled solution to it.

In particular we present a steady-state analysis of the feedback in Tele-Graffiti. Much of the analysis is applicable to other camera-projector systems. Starting with photometric models of the projector, the camera, and the paper used to sketch on, we derive the gain of the system, the final state that the system converges to, and the final image that will be viewed by the users. Based on this result, we derive the "optimal gain" as the gain that results in the final viewed image being as close as possible to the sum of the two original, non-feedback, images. Another common problem with LCD projector-based desktop applications is that the ambient light is sometimes very weak compared to the projector. (This was the problem at the ICCV demo.) We propose a way of using the projector to augment the ambient light in the scene, and derive appropriate settings for this scheme. We also briefly discuss how our analysis is affected by au-



(a) Schematic Diagram of a Two Site Tele-Graffiti System



(b) Image of a Real Tele-Graffiti Site

Figure 1: (a) A schematic diagram of two Tele-Graffiti sites connected by a computer network [Takao *et al.*, 2003]. (b) One site in a real Tele-Graffiti system. The cameras at both sites continuously capture video of what is written on the paper. The video is digitized, compressed, and transmitted to the other site where it is projected onto the paper. Since the pieces of paper can be moved, they have to be tracked in real time at both sites. We can then warp the video appropriately to correctly align the projected video with the paper. See the accompanying movie "telegraffiti-uidemo.mpg" for an example of two people interacting using Tele-Graffiti.

tomatic gain control (AGC) and color imagery.

The final "steady state" images viewed by the users of Tele-Graffiti are a weighted combination of the images that would have been imaged with Tele-Graffiti switched off. In many scenarios it is desirable to estimate what was actually drawn on the paper by the system users. We end this paper by proposing an algorithm to decompose the steady state images into estimates of what was drawn on the paper.

2 Image Formation Model

We illustrate our image formation model in Figure 2. Light is projected by the projector, reflected by the mirror onto the paper where it is combined with the ambient light in the scene, and reflected again to be captured by the camera.

2.1 Projector Model

We assume that if a pixel in the LCD projector is set to project intensity $\text{in}_i^p \in [0, 255]$, the radiance of the light transmitted by the LCD image plane is:

$$l_i^p = g_i^p \cdot \text{in}_i^p \tag{1}$$

where g_i^p is the gain of the projector; i.e. we assume that the response of the projector is linear. The derivation of the relationship between the radiance of the light transmitted by the LCD and the irradiance of the light that is finally received by the paper is similar to the derivation of the relationship between scene radiance and image irradiance for a camera (see [Horn, 1996] Section 10.3) except that the plane of the paper that the light is projected onto is not frontal, unlike the camera image plane. Taking into account this foreshortening, the relationship becomes:

$$l_i^m = l_i^p \cdot \frac{A_i^p}{(d_i^p)^2} \cdot \cos^3 \alpha_i^p \cdot \cos \theta_i^p \tag{2}$$

where l_i^m is the irradiance of the light received at the paper, A_i^p is the area of the projector lens, d_i^p is the distance from the lens to the paper (via the planar mirror, which can be ignored otherwise), α_i^p is the angle that the principal ray makes with the optical axis, and θ_i^p is the angle that the principal ray makes with the normal of the paper.

2.2 Paper Model

We assume that the paper can be modeled as Lambertian. If the total ambient incoming irradiance is l_i^a and the albedo of the paper is al_i, the radiance of the reflected light is:

$$l_i^r = \text{al}_i \cdot [l_i^m + l_i^a]. \tag{3}$$

2.3 Camera Model

The model of the camera is similar to the model of the projector. If the radiance of the paper is l_i^r, the irradiance of the light captured on the image plane of the camera is:

$$l_i^c = l_i^r \cdot \frac{A_i^c}{(f_i^c)^2} \cdot \cos^4 \alpha_i^c \tag{4}$$

where A_i^c is the area of the lens, f_i^c is the distance between the lens and the image plane, and α_i^c is the angle between the principal ray and the optical axis [Horn, 1996]. We also assume that the response of the camera is linear:

$$\text{in}_i^c = g_i^c \cdot l_i^c \tag{5}$$

where g_i^c is the gain of the camera and $\text{in}_i^c \in [0, 255]$.

2.4 Complete Imaging Model

Putting together Equations (1)–(5), we obtain a relationship between the intensity of the pixel in the projector in_i^p and the intensity of the corresponding pixel in the camera in_i^c:

$$\text{in}_i^c = K_i^c \cdot \text{al}_i \cdot [l_i^a + K_i^p \cdot \text{in}_i^p] \tag{6}$$

where $K_i^c = g_i^c \cdot \frac{A_i^c}{(f_i^c)^2} \cdot \cos^4 \alpha_i^c$ and $K_i^p = g_i^p \cdot \frac{A_i^p}{(d_i^p)^2} \cdot \cos^3 \alpha_i^p \cdot \cos \theta_i^p$. Equation (6) says that the image captured by the ith

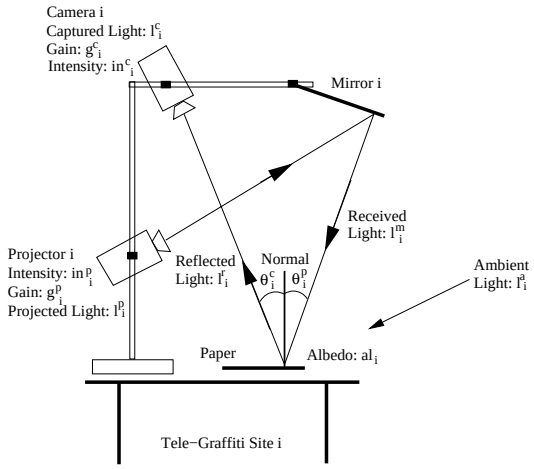


Figure 2: The image formation model. The projector at site i projects a pixel intensity in_i^p with gain g_i^p resulting in projected light radiance l_i^p. The irradiance of the projected light reaching the paper after being reflected by the mirror is l_i^m and the angle that the projected light makes with the paper normal is θ_i^p. The total irradiance of the ambient light that reaches the paper is l_i^a and the albedo of the paper is al_i. The radiance of the reflected light is l_i^r, the angle the reflected light makes with the normal is θ_i^c, the irradiance of the light reaching the camera is l_i^c, the gain of the camera is g_i^c, and the pixel intensity in the camera is in_i^c.

camera in_i^c is the sum of two terms. The first term $K_i^c \cdot \text{al}_i \cdot l_i^a$ is the image of the paper in the ambient light only; it is the image that would have been captured if the projector was not there. The second term $K_i^c \cdot \text{al}_i \cdot K_i^p \cdot \text{in}_i^p$ is the image created by the light originating from the projector.

2.5 Photometric Calibration

The response functions of cameras and projectors are frequently not linear, as assumed in Equations (1) and (5). Empirically we found the response function of the Sony DFW-VL500 cameras that we used to be linear [Takao *et al.*, 2002] (with gamma correction set to "OFF2.") On the other hand, we found the Panasonic PT-L701U projectors that we used to be highly non-linear. We therefore estimated their response function and corrected for the non-linearity in software. See [Takao *et al.*, 2002] for the details.

3 The Final Viewed Images

If we set up a two site Tele-Graffiti system as in Figure 1, the projected image in_i^p is set to equal the image captured at the other site in_{3-i}^c. If either $i = 1$ or $i = 2$, the other site is $3 - i$. If we now suppose time proceeds in a sequence of steps and denote the image captured by the camera in the ith Tele-Graffiti site at time t by $\text{in}_i^c(t)$, Equation (6) becomes:

$$\text{in}_i^c(t) \;=\; \text{in}_i^c(\text{A}) + K_i^c \cdot \text{al}_i \cdot K_i^p \cdot \text{in}_{3-i}^c(t-1) \qquad (7)$$

where $\text{in}_i^c(\text{A}) = K_i^c \cdot \text{al}_i \cdot l_i^a$ is the image of the paper in the ambient light only. Denoting $G_i = K_i^c \cdot \text{al}_i \cdot K_i^p$ and expanding the recursive term $\text{in}_{3-i}^c(t-1)$ yields:

$$\text{in}_i^c(t) \;=\; \text{in}_i^c(\text{A}) + G_i \cdot \text{in}_{3-i}^c(\text{A}) + G_i \cdot G_{3-i} \cdot \text{in}_i^c(t-2). \quad (8)$$

From Equation (6) half of the initial conditions of this recurrence relation are:

$$\text{in}_i^c(0) \;=\; \text{in}_i^c(\text{A}) + G_i \cdot \text{in}_i^p(0) \qquad (9)$$

where $\text{in}_i^p(0)$ is the image projected by the projector in the ith site at system startup. The other half of the initial conditions are: $\text{in}_i^c(1) \;=\; \text{in}_i^c(\text{A}) + G_i \cdot \text{in}_{3-i}^c(0) \;=$

$$\text{in}_i^c(\text{A}) + G_i \cdot \text{in}_{3-i}^c(\text{A}) + G_i \cdot G_{3-i} \text{in}_{3-i}^p(0). \qquad (10)$$

The solution of the recurrence relation in Equation (8) is a pair of fairly complex expressions, one for t even and one for t odd. See [Takao *et al.*, 2002] for the full details. In either case, in the limit $t \to \infty$ the expression simplifies to:

$$\text{in}_i^c(\infty) \;=\; \frac{\text{in}_i^c(\text{A}) + G_i \cdot \text{in}_{3-i}^c(\text{A})}{1 - G_i \cdot G_{3-i}} \qquad (11)$$

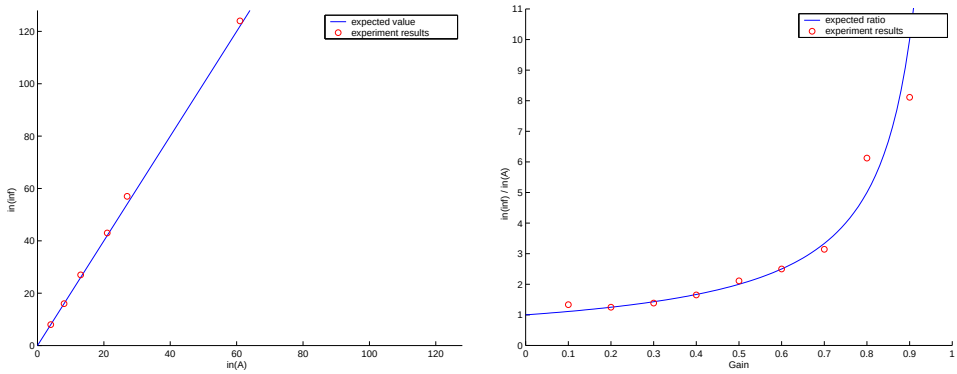
assuming that $G_i \cdot G_{3-i} < 1$. If $G_i \cdot G_{3-i} \geq 1$ the value of $\text{in}_i^c(t)$ keeps increasing until it saturates the camera and the projector and $\text{in}_i^p(t) = \text{in}_i^c(t) = 255$.

Note that the gain $G_i = K_i^c \cdot \text{al}_i \cdot K_i^p$ is not a constant, but depends on the paper albedo. The gain is therefore different for points on the paper that have been written on compared to points that have not been written on. The terms K_i^c and K_i^p also vary across the surface of the paper because the angles α_i^c, α_i^p, and θ_i^p vary, albeit very slowly. Because the variation is so slow, for simplicity we assume that K_i^c and K_i^p are constant in the remainder of this paper.

3.1 Empirical Validation

In our empirical validation we use a blank piece of paper. This is sufficient to validate Equation (11). The albedos al_i and gains G_i are therefore treated as constant in this section. The gain G_i depends on the camera and projector parameters through K_i^c and K_i^p. It is generally easier to change the camera parameters, especially with the Sony DFW-VL500's since their parameters can be changed over the Firewire interface. Note that until Section 6 we assume that the camera Automatic Gain Control is switched off.

The gain G_i is set in the following way. We place a blank piece of paper under the camera and grab an image with the projector switched off. We then set the projector to project intensity $\text{in}_i^p = 128$ and grab a second image. We subtract the first image from the second and estimate the average intensity in a small region in the center of the paper. We then perform a binary search over the camera parameters until this average value equals $128 \times G_i$. (Either the shutter or aperture can be varied. We use the aperture.) Naturally, for each trial with new camera parameters we grab a new image with the projector switched off.



(a) Varying the Intensity $\text{in}_1^c(\text{A}) = \text{in}_2^c(\text{A})$ (b) Varying the Gain $G_1 = G_2$

Figure 3: Empirical validation of Equation (11). (a) We plot $\text{in}_1^c(\infty)$ for various different values of $\text{in}_1^c(\text{A}) = \text{in}_2^c(\text{A})$ for fixed $G_1 = G_2 = 0.5$. (b) We plot the value of $\text{in}_1^c(\infty)/\text{in}_1^c(\text{A})$ for various settings of $G_1 = G_2$ and $\text{in}_1^c(\text{A}) = \text{in}_2^c(\text{A})$. In both graphs we compare the empirically measured results with the values predicted by Equation (11). We find close agreement in both cases validating Equation (11).

Our results are presented in Figure 3. We set the gains $G_1 = G_2$ and intensities $\text{in}_1^c(\text{A}) = \text{in}_2^c(\text{A})$ to be various values and compare the final measured image intensity $\text{in}_i^c(\infty)$ with that predicted by Equation (11). The ambient light image $\text{in}_i^c(\text{A})$ is varied independently of the gain in our experiments by varying the amount of ambient light in the room using an array of spot-lights. In Figure 3(a) we present results varying the ambient light images, keeping the gains fixed. In Figure 3(b) we present results varying the gains, and implicitly the ambient light images also. In Figure 3(b) we therefore plot the ratio of the final viewed image to the ambient light image, a quantity that should be constant if Equation (11) is correct. As can be seen from the graphs in Figure 3, the predicted values closely match those measured empirically, thereby validating Equation (11).

4 Choosing the Optimal Gain

Equation (11) describes the final state that a two site Tele-Graffiti system will settle into. It also describes the final image captured by the cameras and viewed by the users of the system. How can we set the gains G_i and G_{3-i} to maximize the image quality of $\text{in}_i^c(\infty)$?

Equation (11) implies that should we want $\text{in}_i^c(\text{A})$ and $\text{in}_{3-i}^c(\text{A})$ to contribute equal weight to $\text{in}_i^c(\infty)$ we should use a gain $G_i \approx 1$; if $G_i \ll 1$ the user will not see the image $\text{in}_{3-i}^c(\text{A})$ from the other site. On the other hand, in the limit $G_i, G_{3-1} \to 1$ the final viewed image $\text{in}_i^c(\infty) \to \infty$ and the cameras saturate. Too large a gain also leads to "visual echoing." When one of the users' waves their hand across the paper they see multiple "echoes" of it, transmitted back from the image projected at the other site. Another side effect of visual echoing is that, because of inevitable errors in geometric calibration and paper tracking, sketches "run." See Figure 4(a-c) for an example of such visual echoing.

In practice we need to find a compromise between these two extremes. One way to choose the gains is to choose them to make $\text{in}_i^c(\infty)$ as close as possible to $\text{in}_i^c(\text{A}) + \text{in}_{3-i}^c(\text{A})$. It turns out that it is best to use the relative error between these two quantities. If we add in a similar expression for $\text{in}_{3-i}^c(\infty)$, the other site, we aim to minimize:

$$\left(\left[\frac{\text{in}_i^c(\text{A}) + G_i \cdot \text{in}_{3-i}^c(\text{A})}{1 - G_i \cdot G_{3-i}} - \text{in}_i^c(\text{A}) - \text{in}_{3-i}^c(\text{A}) \right]^2 + \left[\frac{\text{in}_{3-i}^c(\text{A}) + G_{3-i} \cdot \text{in}_i^c(\text{A})}{1 - G_i \cdot G_{3-i}} - \text{in}_i^c(\text{A}) - \text{in}_{3-i}^c(\text{A}) \right]^2 \right) / \left[\text{in}_i^c(\text{A}) + \text{in}_{3-i}^c(\text{A}) \right]^2 .$$

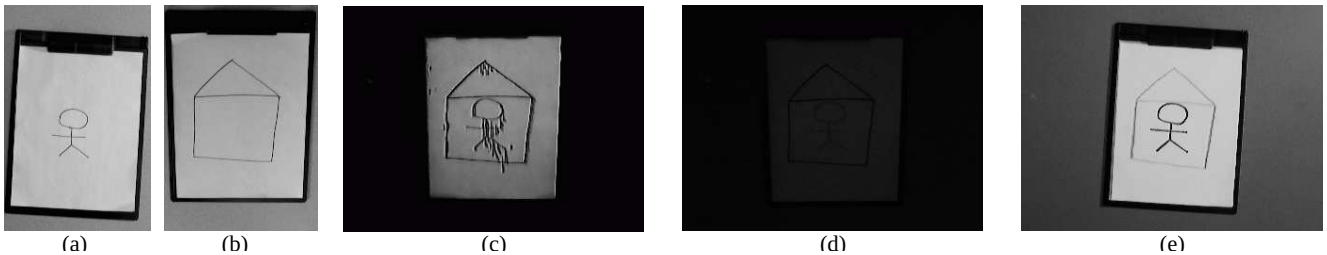
In this expression we have ignored the fact that the gains G_i and G_{3-i} are variables and depend upon the albedo of the paper. Assume that $w(\text{al}_i)$ denotes the probability that the albedo of the paper is al_i. We can then generalize the expression in the equation above to:

$$\int_0^{1/\pi} \left(\left[\frac{\text{in}_i^c(\text{A}) + G_i \cdot \text{in}_{3-i}^c(\text{A})}{1 - G_i \cdot G_{3-i}} - \text{in}_i^c(\text{A}) - \text{in}_{3-i}^c(\text{A}) \right]^2 + \left[\frac{\text{in}_{3-i}^c(\text{A}) + G_{3-i} \cdot \text{in}_i^c(\text{A})}{1 - G_i \cdot G_{3-i}} - \text{in}_i^c(\text{A}) - \text{in}_{3-i}^c(\text{A}) \right]^2 \right) \frac{w(\text{al}_i) w(\text{al}_{3-i})}{\left[\text{in}_i^c(\text{A}) + \text{in}_{3-i}^c(\text{A}) \right]^2} \, \text{d al}_i \, \text{d al}_{3-i}.$$

The ranges of the integrals are $[0, 1/\pi]$ because the maximum value that albedos can take is $1/\pi$ [Horn, 1996]. The equation above contains two types of terms, gains G_i and ambient light images $\text{in}_i^c(\text{A})$. Because $G_i = K_i^c \cdot \text{al}_i \cdot K_i^p$ and $\text{in}_i^c(\text{A}) = K_i^c \cdot \text{al}_i \cdot l_i^a$ these quantities are related by:

$$\text{in}_i^c(\text{A}) \;=\; G_i \cdot \frac{l_i^a}{K_i^p}. \qquad (12)$$

In order to proceed we now make the following three simplifying assumptions: (1) we assume that l_i^a is constant



(a) (b) (c) (d) (e)

Figure 4: (a-c) Visual echoing caused by using too large a gain $G_i = 0.95$, combined with the inevitable errors in tracking the paper and calibrating H_{pc}: (a) $\mathrm{in}_1^c(\mathrm{A})$, (b) $\mathrm{in}_2^c(\mathrm{A})$, (c) $\mathrm{in}_1^c(\infty)$. (d) The final viewed image $\mathrm{in}_1^c(\infty)$ with optimal gain $G_i^{\mathrm{p}} = 0.5$ and under normal office illumination is very dark. (e) The final viewed image $\mathrm{in}_1^c(\infty)$ with optimal gain $G_i^{\mathrm{p}} = 0.5$ and under strong ambient illumination.

across the paper, (2) we assume that K_i^p is constant across the paper, and (3) we assume that $\frac{l_i^a}{K_i^p}$ is the same at the two sites. The first of these three assumptions just means that the ambient light is constant across the paper which is very reasonable. The second assumption just means that the various angles between the projected and reflected light, the paper, the camera the projector are all constant across the paper. This is also a reasonable approximation. The third assumption can easily be generalized, however. The details are omitted but it is straightforward to estimate $\frac{l_i^a}{K_i^p}$ for each of the two sites separately from Equation (12). The constant relative values of these expressions can then included and different optimal gains derived for the two sites.

Even after making these assumptions, the optimal value clearly still depends on the weighting function $w(\mathrm{al}_i)$. The sketches we work with are mostly simple "black on white" line drawings. The albedo is therefore nearly always either $\mathrm{al}_i^{\mathrm{p}}$, the paper albedo, or 0. Let us therefore assume that:

$$w(\mathrm{al}_i) \;=\; \frac{1}{2}\delta(\mathrm{al}_i) + \frac{1}{2}\delta(\mathrm{al}_i - \mathrm{al}_i^{\mathrm{p}}) \qquad (13)$$

where $\delta(\cdot)$ is the Dirac delta function; i.e. a unit impulse. It is possible to continue the analysis with other choices of the weighting function $w(\mathrm{al}_i)$ and derive the optimal gain in other scenarios. We just need to make a concrete choice to proceed. Let us denote $G_i^{\mathrm{p}} = K_i^c \cdot \mathrm{al}_i^{\mathrm{p}} \cdot K_i^p$, the gain of the paper. The optimality criterion then simplifies to:

$$\left(\left[\frac{(G_i^{\mathrm{p}})^2 \cdot G_{3-i}^{\mathrm{p}} + (G_i^{\mathrm{p}} \cdot G_{3-i}^{\mathrm{p}} + G_i^{\mathrm{p}} - 1)\cdot G_{3-i}^{\mathrm{p}}}{1 - G_i^{\mathrm{p}} \cdot G_{3-i}^{\mathrm{p}}}\right]^2 + \left[\frac{G_i^{\mathrm{p}} \cdot (G_{3-i}^{\mathrm{p}})^2 + (G_i^{\mathrm{p}} \cdot G_{3-i}^{\mathrm{p}} + G_{3-i}^{\mathrm{p}} - 1)\cdot G_i^{\mathrm{p}}}{1 - G_i^{\mathrm{p}} \cdot G_{3-i}^{\mathrm{p}}}\right]^2 + (G_i^{\mathrm{p}})^2 + (G_{3-i}^{\mathrm{p}})^2\right) \Big/ \left[G_i^{\mathrm{p}} + G_{3-1}^{\mathrm{p}}\right]^2. \qquad (14)$$

See [Takao *et al.*, 2002] for more details of the derivation. The minimum value of the expression in Equation (14) is easily found numerically to be $G_i^{\mathrm{p}} = G_{3-i}^{\mathrm{p}} = 0.5$. Equation (14) is plotted in Figure 5(a) for $G_i^{\mathrm{p}} = G_{3-i}^{\mathrm{p}}$. The optimal gain of the paper is therefore 0.5. The gain of the paper G_i^{p} can be set using the method described in Section 3.1.

5 Illumination Augmentation

Most LCD projectors are designed to be very bright because of the quadratic falloff in brightness with the distance d_i^p between the projector and the projection surface. See Equation (2). In Tele-Graffiti the value of d_i^p is smaller than the value that most projectors are designed for. As a result K_i^p is often relatively large. To obtain $G_i^{\mathrm{p}} = 0.5$, we have to use a small aperture (or a fast shutter) to make K_i^c small enough. The result is that $\mathrm{in}_i^c(\mathrm{A}) = K_i^c \cdot \mathrm{al}_i \cdot l_i^a$, the image of the paper in ambient light, is very dark. As a result, the final image viewed $\mathrm{in}_i^c(\infty)$ is also very dark. See Figure 4(d) for an example. Ideally we would like $\mathrm{in}_i^c(\infty)$ to have its brightest pixels to have intensity close to $\mathrm{in}_i^{\max} \approx 255$ to maximize the dynamic range of the camera. Assume that the paper is the brightest object. Then, to obtain:

$$\mathrm{in}_i^c(\infty) \;=\; \frac{\mathrm{in}_i^c(\mathrm{A}) + G_i^{\mathrm{p}} \cdot \mathrm{in}_{3-i}^c(\mathrm{A})}{1 - G_i^{\mathrm{p}} \cdot G_{3-i}^{\mathrm{p}}} \;=\; \mathrm{in}_i^{\max} \;\approx\; 255 \qquad (15)$$

when $\mathrm{in}_i^c(\mathrm{A}) = \mathrm{in}_{3-i}^c(\mathrm{A})$ and $G_i^{\mathrm{p}} = G_{3-i}^{\mathrm{p}} = 0.5$, we need:

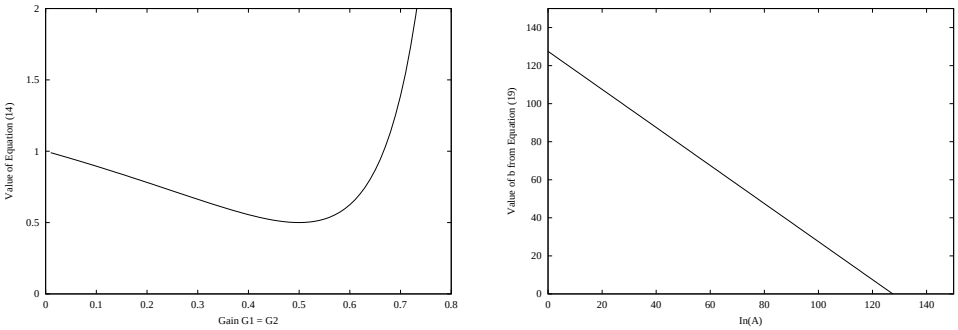
$$\mathrm{in}_i^c(\mathrm{A}) = \frac{1 - (G_i^{\mathrm{p}})^2}{1 + G_i^{\mathrm{p}}} \times \mathrm{in}_i^{\max} = (1 - G_i^{\mathrm{p}}) \times \mathrm{in}_i^{\max} \approx 127.5. \qquad (16)$$

If the brightest pixel in $\mathrm{in}_i^c(\mathrm{A}) \ll 127$, the brightest pixel in $\mathrm{in}_i^c(\infty)$ will be $\ll 255$ and a large fraction of the dynamic range of the camera (and projector) will be wasted.

The only way to increase $\mathrm{in}_i^c(\mathrm{A}) = K_i^c \cdot \mathrm{al}_i \cdot l_i^a$ is to increase the ambient light l_i^a because the albedo al_i is fixed and K_i^c must to be set to ensure that the gain $G_i^{\mathrm{p}} = 0.5$. It is often impossible or inconvenient to increase the ambient illumination in the room. Fortunately, as we now show how, it is possible to use the projector to provide the additional illumination. Suppose that instead of setting the projected image in_i^p to be the captured image in_{3-i}^c we set it to be:

$$\mathrm{in}_i^p \;=\; \mathrm{b} + \frac{\mathrm{in}_i^{\max} - \mathrm{b}}{\mathrm{in}_i^{\max}} \times \mathrm{in}_{3-i}^c; \qquad (17)$$

i.e. we effectively add constant additional illumination of intensity b. In making this change, we effectively change the gain of the projector to be smaller by the factor $(\mathrm{in}_i^{\max} -$



(a) Derivation of the Optimal Gain (b) The Base Intensity b to Project

Figure 5: (a) Choosing the Gain. We choose the optimal gains G_i^{p} and G_{3-i}^{p} to make $\mathrm{in}_i^c(\infty)$ as close as possible to $\mathrm{in}_i^c(\mathrm{A}) + \mathrm{in}_{3-i}^c(\mathrm{A})$; i.e. to minimize Equation (14). The minimum value is achieved when $G_i^{\mathrm{p}} = G_{3-i}^{\mathrm{p}} = 0.5$. Here we plot Equation (14) for $G_i^{\mathrm{p}} = G_{3-i}^{\mathrm{p}}$. (b) Choosing the additional ambient light b to be projected. Given the brightest pixel in $\mathrm{in}_i^c(\mathrm{A})$ and assuming $G_i^{\mathrm{p}} = 0.5$, the graph above (b) of Equation (19) gives the value of b to be projected using Equation (17) to obtain the brightest pixel in $\mathrm{in}_i^c(\infty)$ to be $\mathrm{in}_i^{\max} = 255$.

$\mathrm{b})/\mathrm{in}_i^{\max}$. We therefore have to increase the gain of the camera by the factor $\mathrm{in}_i^{\max}/(\mathrm{in}_i^{\max} - \mathrm{b})$ to keep the gain of the overall system constant. The value of $\mathrm{in}_i^c(\mathrm{A})$ will therefore increase by this factor because of the change in the gain of the camera and by the amount $\mathrm{in}_i^{\max} \cdot G_i^{\mathrm{p}} \cdot \mathrm{b}/(\mathrm{in}_i^{\max} - \mathrm{b})$ because of the additional ambient light. To choose b to optimize the dynamic range of the camera we need to solve:

$$\frac{\mathrm{in}_i^{\max}}{\mathrm{in}_i^{\max} - \mathrm{b}} \times [\mathrm{in}_i^c(\mathrm{A}) + G_i^{\mathrm{p}} \times \mathrm{b}] \;=\; (1 - G_i^{\mathrm{p}}) \times \mathrm{in}_i^{\max} \quad (18)$$

The solution of this equation is:

$$\mathrm{b} \;=\; (1 - G_i^{\mathrm{p}}) \times \mathrm{in}_i^{\max} - \mathrm{in}_i^c(\mathrm{A}). \quad (19)$$

Equation (19), plotted in Figure 5(b) for $G_i^{\mathrm{p}} = 0.5$ and $\mathrm{in}_i^{\max} = 255$, can be used to estimate a suitable base intensity b to project using Equation (17) to ensure that the overall system does not saturate, and yet the maximum possible dynamic range of the camera and projector are used. Example images of the system running using the value of b computed using Equation (19) are included in Figure 6. Note that these images have the image quality that we have been striving for; they are bright, have excellent contrast, and there is minimal visual echoing. Also note that the intensity correction in Equation (17) can easily be implemented as a table lookup. See [Takao *et al.*, 2002].

6 Automatic Gain Control

We have analyzed the Tele-Graffiti feedback loop assuming the cameras are set up with fixed apertures and shutter speeds to give a constant gain, however Tele-Graffiti is normally operated with the automatic gain control (AGC) switched on. How does using AGC affect our analysis?

The analysis itself is not affected. The only thing that is affected is our ability to set the gain independently. Instead, AGC operates by adjusting the aperture (or shutter speed) to enforce the requirement that the average intensity in a certain area of the image (normally the center) is a fixed value, the exposure. When AGC is switched on, the camera gain is adjusted until Equation (11) is satisfied, with $\mathrm{in}_i^c(\infty)$ equal to the exposure. If the ambient light level is low, this results in a large gain $G_i \approx 1$ and the system has substantial visual echoing. If the ambient light level is high, the gain $G_i \approx 0$ and what is written at the other site cannot be seen.

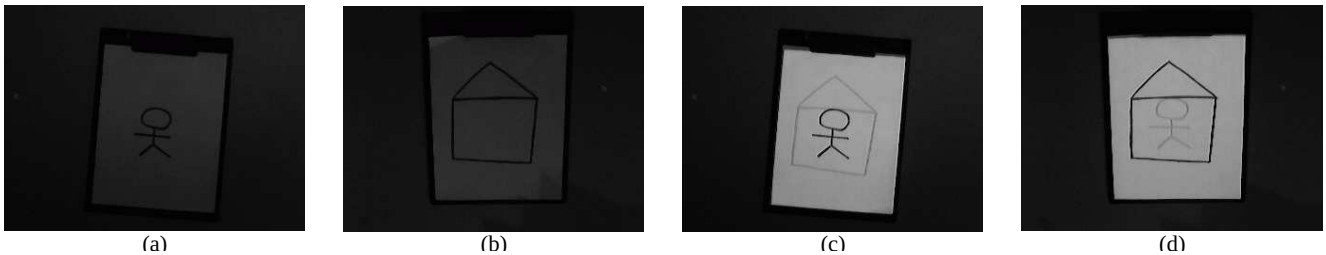
When using automatic gain control, the only parameter we can use to indirectly control the gain is the amount of additional ambient light provided by the projector. In particular, the offset b is the one parameter we can use to change the gain and thereby the image quality. As we increase b the gain of the cameras will be reduced and there will be less visual echoing. On the other hand, as we increase b, the image from the other site will be less visible.

Using a similar procedure to that in Section 3.1 we can build a lookup table for the gain as a function of b. See [Takao *et al.*, 2002] for the details and an example of such a table. We can then use this table to set the gain to be any desired value. The results are shown in Figure 7.

In practice we found that the best way to run Tele-Graffiti is to use AGC. If there is insufficient light there will be visual echoing. The value of b should then be adjusted by hand until the visual echoing stops and the right trade-off between dynamic range and lack of echoing is achieved.

7 Color Imagery

Working with color images simply means that there is a separate gain for each channel. Assuming the camera and projector are reasonably color balanced these gains are fairly similar and the analysis above can be performed with the average of the three gains. If there is substantial difference in the gains, obtaining good image quality can be difficult. If we set up the system so that the average gain of the paper is 0.5, the gain of one of the channels may be much larger; in extreme cases the gain may be ≈ 1.0. In such cases one of the color channels may saturate. An example of such a



(a) (b) (c) (d)

Figure 6: A 2 site Tele-Graffiti system operating with optimal gain $G_i^p = 0.5$ and additional ambient light b provided by the projector: (a) = $\text{in}_1^c(\text{A})$, (b) = $\text{in}_2^c(\text{A})$, (c) = $\text{in}_1^c(\infty)$, (d) = $\text{in}_2^c(\infty)$. Note the superior image quality of the images compared to Figure 4(c-d).

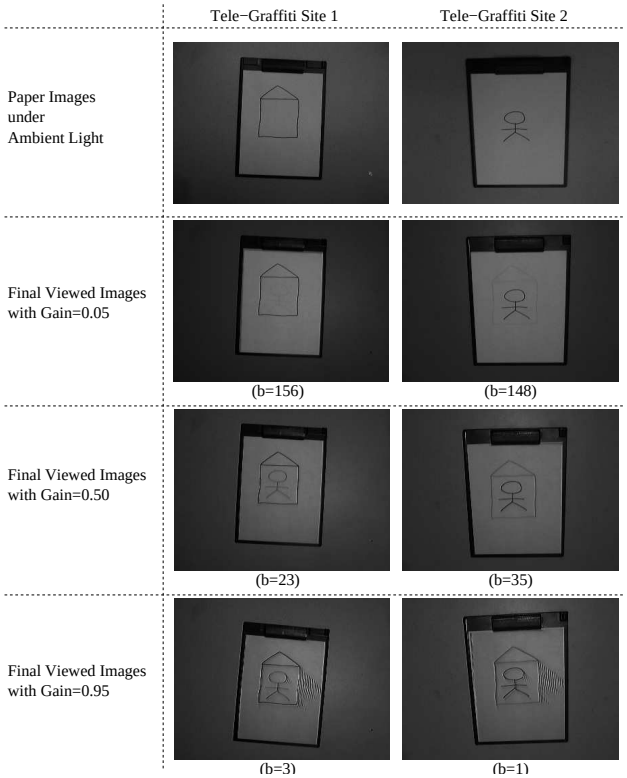


Figure 7: Final viewed images with automatic gain control turned on. The final viewed images ($\text{in}_i^c(\infty)$) with gain=0.05, 0.50, 0.95 are shown in the second, third and fourth rows respectively. To obtain these gains, we computed the base intensity to project b using the lookup table described in [Takao *et al.*, 2002].

situation is included in [Takao *et al.*, 2002].

8 Image Separation

When we switch on Tele-Graffiti the system quickly converges to the steady state:

$$\text{in}_i^c(\infty) \;=\; \frac{\text{in}_i^c(\text{A}) + G_i \cdot \text{in}_{3-i}^c(\text{A})}{1 - G_i \cdot G_{3-i}}. \qquad (20)$$

Is it possible to reverse this process? Can we recover what is drawn on the paper from what is imaged? Doing this corresponds to inverting Equation (20). As in Section 4 the difficulty in doing this is the fact that G_i is not constant, but is related to $\text{in}_i^c(\text{A})$ through:

$$\text{in}_i^c(\text{A}) \;=\; G_i \cdot \frac{l_i^a}{K_i^p}. \qquad (21)$$

If the ambient light l_i^a can vary completely independently across the paper, Equation (20) is not invertible. In Section 4 we assumed that l_i^a and K_i^p are constant across the paper. To invert Equation (20) we now make the same assumption. We also assume that the value of $\frac{l_i^a}{K_i^p}$ has been estimated at system startup. This can easily be performed. We set the gain of the paper G_i^{p} to a fixed value. We can then measure the average ambient light image of the paper and use the results to compute $\frac{l_i^a}{K_i^p}$ from Equation (21).

Once $\frac{K_i^p}{l_i^a}$ is known, it is straightforward to invert Equation (20), a pair of simultaneous quadratic equations. For lack of space, the reader is referred to [Takao *et al.*, 2002] for the details. The result, however are closed form solutions for the ambient light images $\text{in}_i^c(\text{A})$ and $\text{in}_{3-i}^c(\text{A})$ in terms of the final viewed images $\text{in}_i^c(\infty)$ and $\text{in}_{3-i}^c(\infty)$.

We have implemented this solution to the image separation problem, the results of which are shown in Figures 8. We found that where the intensity is roughly constant the original "drawn" image is restored very well. At points with high gradient, however, there are significant artifacts caused by the "defocussing" effect of several components in the system: the camera, the projector, and the image warping. Defocus effects are not incorporated in our image formation model. Potentially a more sophisticated image formation model could be derived. In the meantime we process the results in Figures 8 with a simple algorithm to remove most of the artifacts. We: (1) detect edges in the final viewed images, (2) compare the magnitudes of the edges at the two sites, and (3) smooth the edges at the "weaker" site based on the assumption that the edges originate from the side at which they are stronger. The results of applying this algorithm to the results in Figure 8. Overall the separated images match the actual drawn images reasonable well, (especially considering that the system was effectively set up in Section 4 to make this task as hard as possible.)

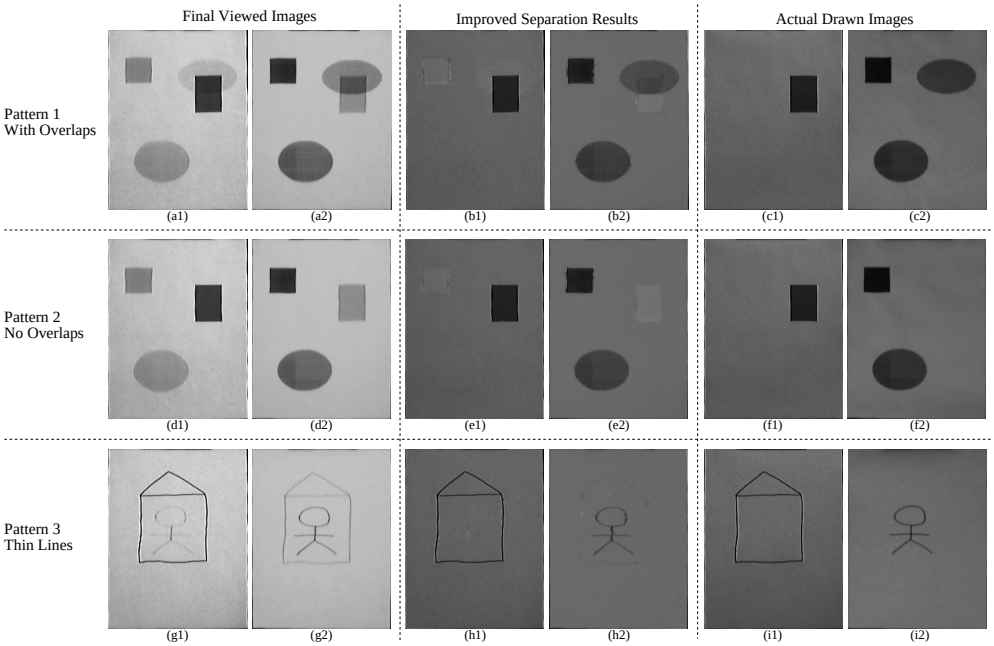


Figure 8: The results of image separation. The first column of each row contains the two final viewed images of the two sites from which we compute the separated images shown in the second column. The third column contains the actual drawn images at the two sites.

9 Discussion

We have performed a steady-state analysis of Tele-Graffiti. We have shown how this analysis can be used to derive the optimal settings for the system and to separate the steady-state images to derive estimates of what is actually drawn on the paper. This steady-state analysis requires a number of assumptions, most notably that what appears on the paper does not change. One possible direction for future work is to perform a "perturbation analysis" to determine how stable the system is when something new is drawn on the paper, or some other change occurs in the system settings.

Acknowledgments

The research described in this paper was conducted while the authors were all at Carnegie Mellon Robotics Institute. We would like to thank Iain Matthews and Bart Nabbe for their help constructing Tele-Graffiti [Takao *et al.*, 2001].

References

[Hall *et al.*, 1999] D. Hall, C. Le Gal, J. Martin, T.Kapuscinski, and J. Crowley. MagicBoard: A contribution to an intelligent office environment. In *Proc. of the 7th Intl. Symp. on Intel. Rob. Sys.*, 1999.

[Horn, 1996] B. Horn. *Robot Vision*. McGraw Hill, 1996.

[Leibe *et al.*, 2000] B. Leibe, T. Starner, W. Ribarsky, Z. Wartell, D. Krum, B. Singletary, and L. Hodges. The perceptive workbench: Towards spontaneous and natural interaction in semi-immersive virtual environments. In *Proc. of the IEEE Virtual Reality Conference*, 2000.

[Raskar and Beardsley, 2001] R. Raskar and P. Beardsley. A self-correcting projector. In *Proc. of CVPR*, 2001.

[Raskar *et al.*, 1998] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *SIGGRAPH*, 1998.

[Sato *et al.*, 2000] Y. Sato, Y. Kobayashi, and H. Koike. Fast tracking of hands and fingertips in infrared images for augmented desk interface. In *Proc. of the 4th Intl. Conf. on Automatic Face and Gesture Recognition*, 2000.

[Sukthankar *et al.*, 2001a] R. Sukthankar, T.-J. Cham, and G. Sukthankar. Dynamic shadow elimination for multi-projector displays. In *Proc. of CVPR*, 2001.

[Sukthankar *et al.*, 2001b] R. Sukthankar, R. Stockton, and M. Mullin. Smarter presentations: Exploiting homography in camera-projector systems. In *ICCV*, 2001.

[Takao *et al.*, 2001] N. Takao, J. Shi, S. Baker, I. Matthews, and B. Nabbe. Tele-Graffiti: A paper-based remote sketching system. In *ICCV*, Demo Session, 2001.

[Takao *et al.*, 2002] N. Takao, J. Shi, and S. Baker. Tele-Grafitti. Technical Report CMU-RI-TR-02-10, Robotics Institute, Carnegie Mellon University, March 2002.

[Takao *et al.*, 2003] N. Takao, J. Shi, and S. Baker. Tele-graffiti: A camera-projector based remote sketching system with hand-based user interface and automatic session summarization. *International Journal of Computer Vision*, 53(2):115 – 133, July 2003.

[Wellner, 1993] P. Wellner. Interacting with paper on the DigitalDesk. *Comm. of the ACM*, 36(7):86–96, 1993.