MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28-30, 2000, The University of Tokyo, Japan

13—28
Optimal Grid Pattern for Automated Matching Using Cross Ratio

Chikara Matsunaga* and Kenichi Kanatani†
*Broadcast Division, FOR-A Co. Ltd.
†Department of Computer Science, Gunma University

Abstract

We design an optimal grid pattern such that an observed image of a small portion of it can be matched to its corresponding position in the pattern easily. The grid shape is so determined that the *cross ratio* of adjacent intervals is different everywhere. The cross ratios are generated by an optimal Markov process that maximizes the accuracy of matching. Finally, we show a virtual studio application of our pattern.

1. Introduction

Camera calibration is a first step in all vision and media applications. The standard method for calibration is to place a planar grid pattern of a known geometry and compute the camera parameters by observing its images [5, 6, 7]. For this computation, we must match observed grid points to their positions in the original pattern.

This paper presents a method for automating this process by designing a grid pattern in such a way that the *cross ratio* of adjacent intervals is different everywhere. Since the cross ratio is invariant to perspective projection, observed grid points can be matched to their corresponding positions easily by comparing the cross ratios.

We then *optimize* the grid shape so that the accuracy of matching is maximized in the presence of noise. Introducing a statistical model of image noise, we generate the grid intervals by an optimal Markov process. Finally, we show a virtual studio application [1] of our pattern.

2. Cross Ratio

The cross ratio can be defined in many different ways [3]. Here, we define the cross ratio τ_i of four numbers $\{x_{i-1}, x_i, x_{i+1}, x_{i+2}\}$ by

$$\tau_i = \frac{x_i - x_{i-1}}{x_{i+1} - x_{i-1}} \Bigg/ \frac{x_{i+2} - x_i}{x_{i+2} - x_{i+1}} = \frac{1}{(1 + l_i/l_{i+1})(1 + l_i/l_{i-1})}, \tag{1}$$

where we have defined the ith interval width l_i by

$$l_i = x_{i+1} - x_i. \tag{2}$$

*Address: 2-3-3 Ohsaku, Sakura, Chiba 285-0802 Japan. E-mail: matsunaga@for-a.co.jp
†Address: Kiryu, Gunma 376-8515 Japan. E-mail: kanatani@cs.gunma-u.ac.jp

Our task is to generate a sequence $\{x_i\}$ in such a way that we can easily find the number i for which $\{x_{i-1}, x_i, x_{i+1}, x_{i+2}\}$ have a specified cross ratio τ_i.

Since intervals of very small separation cannot be discerned in the camera image, the ratio of the minimum width $l_{\min}$ to the average interval width l_0 must be specified. We assume that it is input by the user. Since the absolute scale of the pattern does not have any meaning because we analyze camera images of the pattern, we can normalize the average interval width l_0 to be 1 without losing generality.

Our strategy here is to generate not the sequence $\{x_i\}$ directly but the sequence $\{\tau_i\}$ of cross ratios by a *stochastic process*. The sequence $\{x_i\}$ is easily determined once the cross ratio sequence $\{\tau_i\}$ is given. Suppose we have already generated $\{x_0, ..., x_i, x_{i+1}\}$. If τ_i is given, the next number x_{i+2} is determined from eq. (1) in the form

$$x_{i+2} = x_{i+1} + \frac{\tau_i(1+\gamma_i)}{1 - \tau_i(1+\gamma_i)} l_i, \tag{3}$$

where we have defined the ith *adjacency ratio* γ_i by

$$\gamma_i = \frac{x_{i+1} - x_i}{x_i - x_{i-1}}. \tag{4}$$

Let $p_{l_i,\gamma_i}(\tau)$ be the probability density of the cross ratio τ conditioned on l_i and γ_i defined over a domain $[\tau_{ai}, \tau_{bi}]$. From eq. (3), the condition that the expected length of $x_{i+2} - x_{i+1}$ be 1 is written as

$$\int_{\tau_{ai}}^{\tau_{bi}} \frac{\tau(1+\gamma_i)l_i}{1 - \tau(1+\gamma_i)} p_{l_i,\gamma_i}(\tau) d\tau = 1. \tag{5}$$

3. Error Analysis

Suppose x_{i-1}, x_i, x_{i+1}, and x_{i+2} have errors Δx_{i-1}, Δx_i, Δx_{i+1}, and Δx_{i+2}, respectively. The errors in the intervals l_{i-1}, l_i, and l_{i+1} are

$$\Delta l_{i-1} = \Delta x_i - \Delta x_{i-1}, \quad \Delta l_i = \Delta x_{i+1} - \Delta x_i, \quad \Delta l_{i+1} = \Delta x_{i+2} - \Delta x_{i+1}. \tag{6}$$

It follows from eq. (1) that the cross ratio τ_i has the following error to a first approximation:

$$\Delta\tau_i = \frac{\partial \tau_i}{\partial l_{i-1}}\Delta l_{i-1} + \frac{\partial \tau_i}{\partial l_i}\Delta l_i + \frac{\partial \tau_i}{\partial l_{i+1}}\Delta l_{i+1}$$

$$= -\frac{\partial\tau_i}{\partial l_{i-1}}\Delta x_{i-1} + \left(\frac{\partial\tau_i}{\partial l_{i-1}} - \frac{\partial\tau_i}{\partial l_i}\right)\Delta x_i + \left(\frac{\partial\tau_i}{\partial l_i} - \frac{\partial\tau_i}{\partial l_{i+1}}\right)\Delta x_{i+1} + \frac{\partial\tau_i}{\partial l_{i+1}}\Delta x_{i+2}. \quad (7)$$

If the noise is an independent Gaussian variable of mean 0 and standard deviation σ, the variance of $\Delta\tau_i$ is

$$V[\tau_i] = \sigma^2\left(\left(\frac{\partial\tau_i}{\partial l_{i-1}}\right)^2 + \left(\frac{\partial\tau_i}{\partial l_{i-1}} - \frac{\partial\tau_i}{\partial l_i}\right)^2 + \left(\frac{\partial\tau_i}{\partial l_i} - \frac{\partial\tau_i}{\partial l_{i+1}}\right)^2 + \left(\frac{\partial\tau_i}{\partial l_{i+1}}\right)^2\right), \quad (8)$$

where

$$\frac{\partial\tau_i}{\partial l_{i-1}} = \frac{\tau_i l_i}{l_{i-1}(l_{i-1}+l_i)},$$
$$\frac{\partial\tau_i}{\partial l_i} = -\frac{\tau_i^2(l_{i-1}/l_i + 2 + l_{i+1}/l_i)}{(l_{i-1}/l_i)(l_{i+1}/l_i)}\frac{1}{l_i},$$
$$\frac{\partial\tau_i}{\partial l_{i+1}} = \frac{\tau_i}{l_{i+1}/l_i(1 + l_{i+1}/l_i)}\frac{1}{l_i}. \quad (9)$$

Hence, the standard deviation $\sqrt{V[\tau_i]}$ of τ_i has the following expression:

$$\sqrt{V[\tau_i]} = \frac{\sigma}{l_i}s_{\gamma_i}(\tau_i),$$
$$s_{\gamma_i}(\tau_i) = \sqrt{A_i^2 + (A_i - B_i)^2 + (B_i - C_i)^2 + C_i^2},$$
$$A_i = \frac{\gamma_i^2\tau_i}{1+\gamma_i}, \quad B_i = -\frac{\tau_i^2(1+\gamma_i(2+D_i))}{D_i},$$
$$C_i = \frac{\tau_i}{D_i(1+D_i)}, \quad D_i = \frac{(1+\gamma_i)\tau_i}{1-(1+\gamma_i)\tau_i}. \quad (10)$$

4. Optimal Conditional Probability

If l_i and γ_i are given, the standard deviation $\sqrt{V[\tau_i]}$ is a function of τ_i. It follows that the matching error is minimized if we generate the cross ratio τ densely in the domain over which the standard deviation is small and sparsely in the domain over which it is large. This means that we should define the probability density $p_{l,\gamma}(\tau)$ conditioned on l and γ to be inversely proportional to $s_\gamma(\tau)$, i.e.,

$$p_{l,\gamma}(\tau) = \frac{C_{l,\gamma}}{s_\gamma(\tau)}, \quad (11)$$

where $C_{l,\gamma}$ is the normalization constant. From the normalization condition $\int_{\tau_a}^{\tau_b} p_{l,\gamma}(\tau)d\tau = 1$, we obtain

$$C_{l,\gamma} = 1\Big/\int_{\tau_a}^{\tau_b}\frac{d\tau}{s_\gamma(\tau)}. \quad (12)$$

Substituting eqs. (11) and (12) into eq. (5), we obtain

$$\frac{1+1/l}{1+\gamma}\int_{\tau_a}^{\tau_b}\frac{(1+\gamma)\tau - 1/(1+l)}{1-(1+\gamma)\tau}\frac{d\tau}{s_\gamma(\tau)} = 0. \quad (13)$$

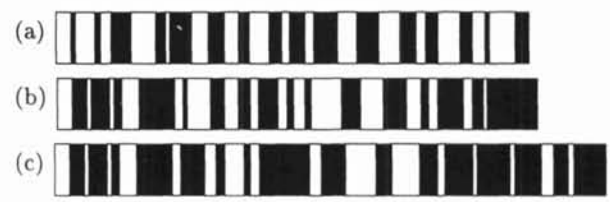

Figure 1: (a) A random sequence. (b) An optimal sequence. (c) An optimal sequence with buffer zones.

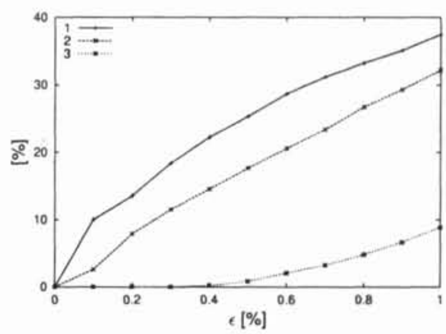

Figure 2: The error ratio of matching. 1. Random sequence. 2. Optimal sequence. 3. Optimal sequence with buffer zones.

It follows that if we define

$$f_{l,\gamma}(x) = \int_{\tau_a}^{x}\frac{\tau - 1/(1+l)(1+\gamma)}{\tau - 1/(1+\gamma)}\frac{d\tau}{s_\gamma(\tau)}, \quad (14)$$

the upper bound τ_b of the domain $[\tau_a, \tau_b]$ is determined for a given lower bound τ_a as the solution of the equation $f_{l,\gamma}(x) = 0$. It has two solutions, one of which is τ_a itself. We denote the other solution by $\tau_{l,\gamma}(\tau_a)$.

5. Optimal Sequence

Measurement error in one position affects two consecutive interval widths, three consecutive adjacency ratios, and four consecutive cross ratios. It follows that the desired distribution of the cross ratio τ depends on the interval width l and the adjacency ratio γ defined by the preceding positions. Hence, the resulting sequence $\{x_i\}$ is a Markov process.

Given x_0, ..., x_i, x_{i+1}, we generate the ith cross ratio τ_i according to the conditional probability (11) over the domain $[\tau_{ai}, \tau_{bi}]$, which is determined from eqs. (1) and (13) in the form

$$\tau_{ai} = \frac{1}{(1+l_i/l_{\min})(1+\gamma_i)}, \quad \tau_{bi} = \tau_{l_i,\gamma_i}(\tau_{ai}). \quad (15)$$

Recall that τ_{bi} is the solution of $f_{l_i,\gamma_i}(x) = 0$ such that $x \neq \tau_{ai}$. The integral in eq. (14) can be numerically evaluated (say, by the trapezoidal rule), and the solution can be obtained by a numerical scheme (e.g., Newton iterations). Then, x_{i+2} is computed by eq. (3), and we repeat this procedure. Let us call the resulting sequence $\{x_i\}$ the *optimal sequence* for short.

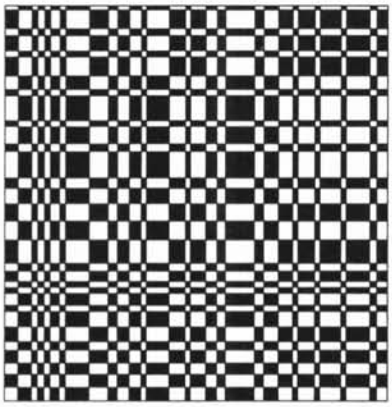

Figure 3: An optimal grid pattern (checkerboard type).

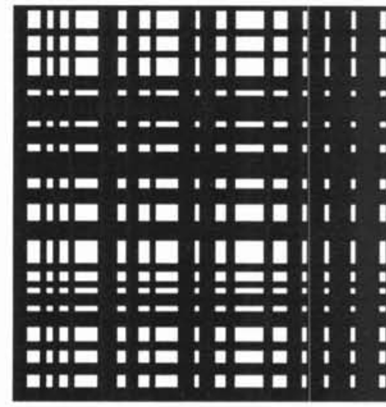

Figure 4: An optimal grid pattern (framework type).

Figure 5: The error ratio of matching. 1. The simple method. 2. Maximum likelihood estimation without using coloring information. 3. Maximum likelihood estimation using checkerboard coloring information. 4. Maximum likelihood estimation using framework coloring information.

Even if the probability distribution of the cross ratio is optimally defined, the resulting sequence may still contain values that are very close to each other as long as the generation is stochastic. This causes deterioration of the matching capability in the presence of image noise. So, we introduce a constraint that no two cross ratios be very close to each other.

If the standard deviation of the noise in $\{x_i\}$ is σ, the standard deviation $s_{l,\gamma}(\tau)$ of τ conditioned on l and γ is given from eqs. (10) in the form

$$s_{l,\gamma}(\tau) = \frac{\sigma}{l} s_\gamma(\tau). \tag{16}$$

Each time we generate a cross ratio τ, we define buffer zones of width $s_{l,\gamma}(\tau)$ on both sides of τ and forbid subsequent cross ratios to occur in those zones.

For comparison, we generate the ith interval l_i independently and uniformly over a domain $[l_{\min}, l_{\max}]$ centered at 1 (i.e., $l_{\max} = 2 - l_{\min}$) and define $x_{i+1} = x_i + l_i$. Let us call the resulting independent additive process $\{x_i\}$ the *random sequence* for short.

Figs. 1(a) and (b) show an instance of the random sequence and the optimal sequence, respectively, for $l_{\min} = 1/4$. Fig. 1(c) is an instance of the optimal sequence with buffer zones, where σ is set to 1% of the average interval width. We added independent Gaussian noise of mean 0 and standard deviation ϵ% of the average interval width to each position and computed the cross ratios of all four consecutive positions. Each position is matched with the position that has the closest cross ratio. Using different noise each time, we repeated this 100 times and plotted the average error ratio for ϵ in Fig. 2.

6. Optimal Grid Pattern

Generating two sequences $\{x_i\}$ and $\{y_j\}$ independently, we can define a grid pattern with vertices $\{(x_i, y_j)\}$. Fig. 3 shows one example for $l_{\min} = 1/3$ with σ set to 1% of the average interval. It is painted like a checkerboard with dark and light blue colors for the convenience of image processing (chromakey application).

In order to compute the cross ratios in both directions, we need to observe at least a 3×3 block, from which the cross ratios are computed in four ways. Let τ_x and τ_y be the averages of the four values for the x and y directions, respectively. The absolute position of that block in the pattern is determined by finding integers i and j such that $|\tau_x - \tau_{x(i)}|$ and $|\tau_y - \tau_{y(j)}|$ are minimized, where $\{\tau_{x(i)}\}$ and $\{\tau_{y(j)}\}$ are the cross ratio sequences associated with $\{x_i\}$ and $\{y_j\}$, respectively.

However, this process does not take into account the error behavior of the cross ratio. Since the standard deviation of the cross ratio can be evaluated by eqs. (10), a statistically optimal method is the *maximum likelihood estimation*: we minimize the squared Mahalanobis distance

$$J(i,j) = \frac{l_{x(i)}^2 |\tau_x - \tau_{x(i)}|^2}{s_{\gamma_{x(i)}}(\tau_{x(i)})^2} + \frac{l_{y(j)}^2 |\tau_y - \tau_{y(j)}|^2}{s_{\gamma_{y(j)}}(\tau_{y(j)})^2}, \tag{17}$$

where $\{l_{x(i)}\}$ and $\{l_{y(j)}\}$ are the interval sequences defined from $\{x_i\}$ and $\{y_j\}$, and $\{\gamma_{x(i)}\}$ and $\{\gamma_{y(j)}\}$ are the similarly defined adjacency ratio sequences.

For a checkerboard pattern, a 3×3 block has two

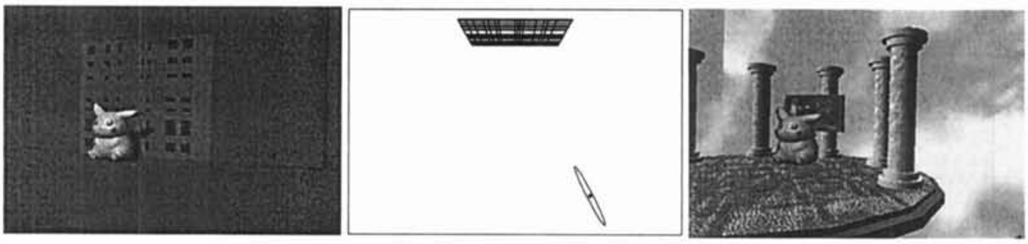

Figure 6: (a) Original image. (b) Estimated camera position and its reliability. (c) A virtual scene generated from (a).

possibilities for its coloring. This information can be used to reduce the search space for minimizing eq. (17). Another possibility for coloring the pattern is to alternate colors for neighboring rows and columns (Fig. 4). Let us call it a *framework pattern*. It has four possibilities for coloring a 3×3 block, reducing the search space to a half that for the checkerboard pattern.

7. Simulation

We added Gaussian random noise of mean 0 and standard deviation ϵ% of the average interval width to the coordinates of each grid point of the pattern shown in Figs. 3 and 4. The positions of all 3×3 blocks are computed from the observed cross ratios, and this process was repeated 100 times using different noise each time. Fig. 5 plots the average error ratio for ϵ. Here, we compared the following four methods:

1. The simple method: $|\tau_x - \tau_{x(i)}|$ and $|\tau_y - \tau_{y(j)}|$ are minimized.
2. Maximum likelihood estimation: eq. (17) is minimized.
3. Maximum likelihood estimation combined with checkerboard coloring information.
4. Maximum likelihood estimation combined with framework coloring information.

We can see that maximum likelihood estimation slightly reduces the matching error as compared with the simple method. In contrast, the coloring information dramatically improves the accuracy and that the framework pattern is more effective than the checkerboard pattern.

8. Virtual Studio Application

Fig. 6(a) is a real image of a toy, behind which we placed our optimal grid pattern. After segmenting the toy image from the background by using a chromakey technique, we computed the 3-D position and focal length of the camera by observing an unoccluded portion of the grid pattern [2, 4]. The focal length is estimated to be 576 pixels. The standard deviations of the focal length, the translation, and the rotation are evaluated to be ± 38.3 pixels, ± 5.73cm, and $\pm 0.812°$, respectively.

Fig. 6(b) is the top view of the estimated camera position and its uncertainty ellipsoid (three times the standard deviation in each orientation). Fig. 6(c) is a composition of the toy image and a graphics scene generated by VRML.

8. Concluding Remarks

We have designed an optimal grid pattern such that an observed image of a small portion of it can be matched to its corresponding position in the pattern easily. The grid shape is so determined that the cross ratio of adjacent intervals is different everywhere. By statistical analysis of image noise, we have generated the cross ratios by an optimal Markov process that maximizes the accuracy of matching.

Acknowledgments: This work was in part supported by the Ministry of Education, Science, Sports and Culture, Japan under a Grant in Aid for Scientific Research C(2) (No. 11680377). The authors thank Yasushi Kanazawa of Toyohashi University of Technology for collaboration in this research.

References

[1] S. Gibbs, C. Arapis, C. Breiteneder, V. Lalioti, S. Mostafawy and J. Speier, Virtual studios: An overview, *IEEE Multimedia*, **5**-1 (1998), 24–35.

[2] C. Matsunaga and K. Kanatani, Calibration of a moving camera using a planar pattern: Optimal computation, reliability evaluation and stabilization by model selection, *Proc. 6th Euro. Conf. Computer Vision*, June 2000, Dublin, Ireland, Vo. 2, pp. 595–609.

[3] J. G. Semple and G. T. Kneebone, *Algebraic Projective Geometry*, Clarendon, Oxford, 1998.

[4] Y. Seo, M.-H. Ahn and K.-S. Hong, A multiple view approach for auto-calibration of a rotating and zooming camera, *IEICE Trans. Inf. & Syst.*, **E83-D**-7 (2000), 1375–1385.

[5] B. Triggs, Autocalibration from planar scenes, *Proc. 5th Euro. Conf. Computer Vision*, June 1998, Freiburg, Germany, Vol. 1, pp. 89–105.

[6] R. Y. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision methodology using off-the-shelf TV cameras and lenses, *IEEE J. Robotics Automation*, **3**-4 (1987), 323–344.

[7] Z. Zhang, Flexible camera calibration by viewing a plane from unknown orientations, *Proc. 7th Int. Conf. Computer Vision*, September 1999, Kerkya, Greece, pp. 666–673.