DYNAMIC HISTOGRAM WARPING OF IMAGE PAIRS FOR CONSTANT IMAGE BRIGHTNESS

Ingemar J. Cox, Sébastien Roy

NEC Research Institute,
4 Independence Way
Princeton, NJ 08540, U.S.A.
ingemar|sebastien@research.nj.nec.com

Sunita L. Hingorani

AT&T Bell Laboratories
184 Liberty Corner Road
Warren, NJ 07059
sunita@cartoon.lc.att.com

ABSTRACT

The constant image brightness (CIB) assumption assumes that the intensities of corresponding points in two images are equal. This assumption is central to much of computer vision. However, surprisingly little work has been performed to support this assumption, despite the fact the many of algorithms are very sensitive to deviations from CIB.

An examination of the images contained in the SRI JISCT stereo database revealed that the constant image brightness assumption is indeed often false. Moreover, the simple additive/multiplicative models of the form $I_L = \beta I_R + \alpha$ do not adequately represent the observed deviations. A comprehensive physical model of the observed deviations is difficult to develop. However, many potential sources of deviations can be represented by a non-linear monotonically increasing relationship between intensities. Under these conditions, we believe that an expansion/contraction matching of the intensity histograms represents the best method to both measure the degree of validity of the CIB assumption and correct for it. Dynamic histogram warping (DHW) is closely related to histogram specification. However, it is shown that histogram specification introduces artifacts that do not occur with dynamic histogram warping.

Experimental results show that image histograms are closely matched after DHW, especially when both histograms are modified simultaneously. DHW is also capable of removing simple constant additive and multiplicative biases without derivative operations, thereby avoiding amplification of high frequency noise. It is demonstrated that DHW can improve the estimates from stereo and optical flow estimators.

1. INTRODUCTION

The constant image brightness (CIB) assumption assumes that the intensities of corresponding points (or planar patches) in two (or more) images are equal. This assumption is central to bodies of work in optical flow estimation, motion and structure, stereo and recognition based on color histograms. However, surprisingly little work has been performed to support this assumption, despite the fact that many of these algorithms are very sensitive to deviations from CIB. We examined 49 image pairs contained in the SRI JISCT stereo database by comparing their intensity histograms. We found that corresponding pairs of histograms could vary significantly, i.e. the constant image brightness assumption is often false. In practice, it is common to believe that any deviation from the constant image brightness assumption can be modelled by a simple global spatially-invariant additive constant and/or a global spatially-invariant scaling of the image intensities (contrast), i.e. $I_A = \beta I_B + \alpha$. However, our experiments show that this linear model does not adequately represent the observed deviations. Experiments suggest that the form of the relationship between the two sets of intensities might be a non-linear model of the form $I_A = \beta I_B^{\gamma} + \alpha$.

There are a number of possible reasons why a pair of images might deviate from the CIB assumption, assuming that the image content remains the same. These include (1) variations in illumination, (2) variations in camera signal response and (3) the time-varying non-linear automatic gain control of the cameras. If it is assumed that these factors can be lumped together and represented as an arbitrary non-linear monotonically increasing function that uniquely maps intensity values in image A to intensity values in image B, then errors in the constant image brightness assumption can be corrected, or at least reduced, by matching the intensity histograms of the two images.

Image histogram matching is typically performed by the process of histogram specification [3]. However, while histogram specification produces good matches, the *local comparison* of histograms introduces artifacts (spikes in the matched histograms) because matching errors propagate and accumulate and must periodically be corrected. The simple example of Figure (1) illus-

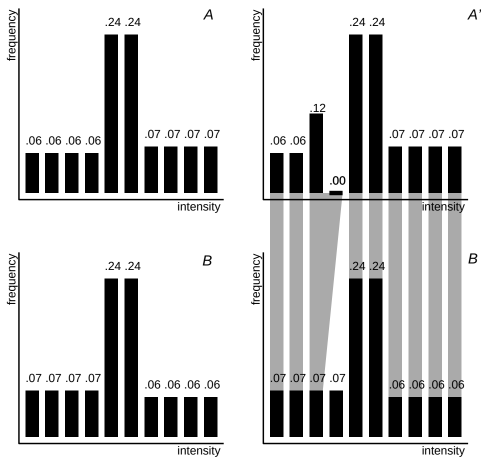

Figure 1: Incorrect histogram specification. A) Original histogram. B) Specified histogram. A') Resulting histogram using SML histogram specification as in [8]. Using GML also gives erroneous results for this case.

trates the effect. Here, intensity values 1 through 4 occur more frequently in image A while intensity values 7 through 10 occur more frequently in image B. Clearly though, the mapping should be one to one, i.e. $I_i^A = I_i^B$. However, the matching of $I_{1,2}^A$ with $I_{1,2}^B$ results in a cumulative error of 0.02, which is subsequently reduced by matching $I_{3,4}^A$ to I_4^B, as shown in histograms A',B of Figure (1). By matching histogram values directly and performing a *global* optimization via dynamic programming, this problem is avoided and better matching is thereby achieved as shown in Figure (1d). To avoid this problem and others reported in [7, 8] we have developed a new approach called dynamic histogram warping (DHW) that uses dynamic programming on the histograms values directly. Histogram specification typically maps an image histogram to a (fixed) reference histogram. The precision of this mapping is limited since while compression (i.e. a many to one mapping) is straightforward, expansion (i.e. mapping a single intensity to many) is not possible.[1]. However, an expansion of one histogram is equivalent to a corresponding compression of the other histogram. Thus, if the reference histogram is also modified by corresponding compressions, a new pair of histograms can be generated that are more closely matched than original histogram specification can achieve. This is important for purposes of constant image brightness correction.

[1]Note that expansion is different from stretching which is a one-to-one mapping in which only the range of intensities is altered.

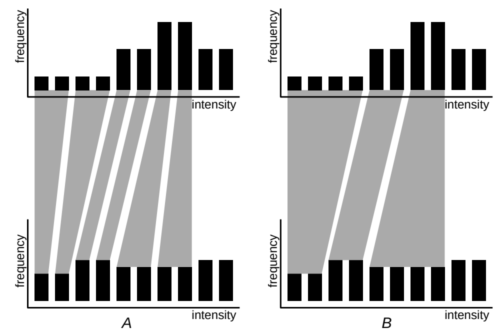

Figure 2: Histogram matching. Legal matches (A) always join one intensity to one or more others. Illegal matches (B) join many intensities to many others.

2. DYNAMIC HISTOGRAM WARPING

Dynamic histogram warping is strongly related to work in sequence comparison and especially to dynamic time warping [5]. For dynamic histogram warping, two intensity histograms are compressed and/or expanded to best match one another. Histogram samples can be matched one-to-one, one-to-many (expansion) or many-to-one (contraction), as illustrated in Figure (2a). However, the many-to-many mappings of Figure (2b) are considered illegal. Because of quantization error, we initially considered allowing many-to-many mappings. However, while the differences in the resulting histograms were reduced with such mappings, the original shape of the histograms was often lost. We felt that it was desirable to retain the original shape as much as possible and therefore did not allow many-to-many mappings, though to do so is straightforward within our framework.[2]

To specify the cost of a matching, let h_m^A and h_n^B represent the frequency of occurrence of the mth and nth intensity values in images A and B respectively. Let H_m^A and H_n^B represent the cumulative frequency of occurence such that $H_m^A = \sum_{i=1}^m h_i^A$ and $H_n^B = \sum_{i=1}^n h_i^B$. Then the usual cost of matching intensity I_m^A of image A with intensity I_n^B in image B is simply $\left|h_m^A - h_n^B\right|$. This is appropriate for a one-to-one mapping. However, for histograms the quantities being compared are the number of *occurrences* of intensity values. Thus, for a one-to-two mapping, for example, the cost should be $\left|h_m^A - (h_n^B + h_{n-1}^B)\right|$ and for a one-to-k mapping $\left|h_m^A - \sum_{i=0}^{k-1} h_{n-i}^B\right|$. The fact that the cost of matching h_{m+1}^A to h_n^B depends on whether or not h_m^A was matched to h_n^B, complicates the dynamic

[2]Of course, an additional cost must be associated with many-to-many mappings otherwise the degenerative mapping of all-to-all is always the optimum solution.

programming. However, since the maximum size of a compression or expansion is always finite[3], then such a cost function can be accomodated [2]. In general, the cost of a k-to-l mapping is

$$\begin{aligned} d_{k,l}(m,n) &= \left| \sum_{i=0}^{k-1} h^A_{m-i} - \sum_{j=0}^{l-1} h^B_{n-j} \right| \\ &= \left| (H^A_m - H^A_{m-k}) - (H^B_n - H^B_{n-l}) \right| \end{aligned}$$

Finally then, it is necessary to define the total cost of a matching. This cost is defined recursively as

$$\begin{aligned} &D(0,0) = 0 \\ &D(i,j) = \infty \qquad (i \le 0, j \le 0, (i,j) \ne (0,0)) \\ &D(m,n) = \min \begin{cases} D(m-1,n-1) + d_{1,1}(m,n) \\ D(m-k,n-1) + d_{k,1}(m,n), (2 \le k \le M) \\ D(m-1,n-l) + d_{1,l}(m,n), (2 \le l \le N) \end{cases} \end{aligned}$$

where M and N represent the maximum allowable compression of the respective histograms. The cost function can be efficiently minimized via dynamic programming.

3. EXPERIMENTAL RESULTS

We applied dynamic histogram warping to all of the SRI JISCT database and compared it with histogram specification. To illustrate its effectiveness, the intensities of the left image of the synthetic stereo pair of Figure (3) were non-linearly mapped using $\lfloor 128 \sin(\frac{I_i}{255}\pi - \frac{\pi}{2}) + 128 \rfloor$ as the transform. The intensity histograms of the two images are shown in Figure (4). Histogram specification and dynamic histogram warping were then applied, using the second histogram as a reference, to produce the corrected histograms of Figure (5). Superior matching is achieved using DHW (sum of absolute differences of 214) while the GML histogram specification yields an error of 222. Notice the large spikes around intensity 100 in Figure (5a), which are artifacts due to the local optimization of the cumulative histograms. Finally, in order to correct for CIB, both histograms were simultaneously corrected using DHW.

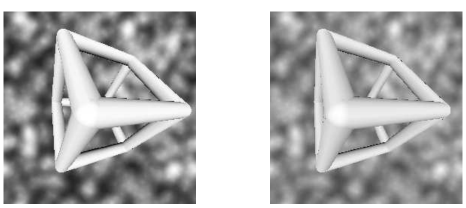

Figure 3: A synthetic stereo pair.

[3] In the limit one-to-N, where N is the range of intensity values.

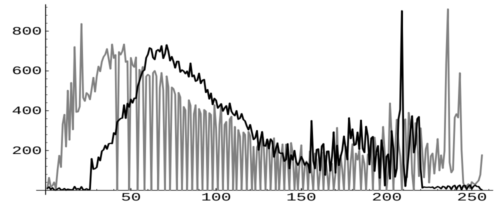

Figure 4: Intensity histograms for the stereo pair. The left and right histograms are respectively in gray and black.

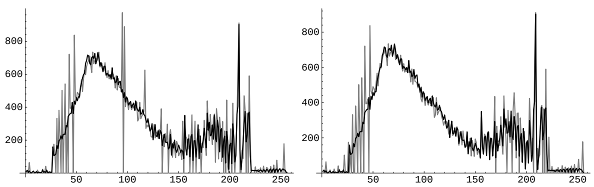

Figure 5: Corrected histograms: (A) histogram specification (GML), (B) DHW with right histogram fixed. The left and right histograms are respectively grey and black.

The corrected histograms are almost identical (error of 24). The utility of this is clearly demonstrated by applying a stereo algorithm [1] that assumes CIB to the image pair (1) with no correction, (2) with correction by histogram specification and (3) correction by DHW. The corresponding disparity maps are shown in Figures (6). Very poor disparity estimates are obtained with no correction. Histogram specification provides substantial improvement but significant errors remain. In contrast, with DHW, the disparity map is almost as good as the algorithm can perform on perfect data.

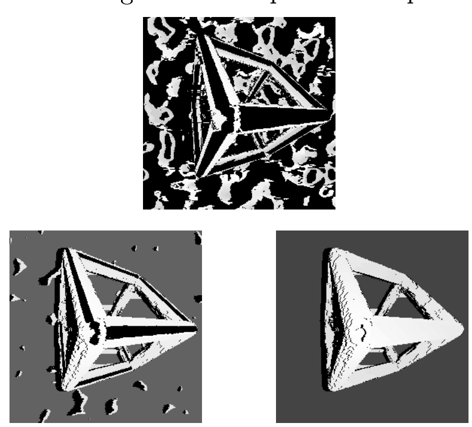

Figure 6: Disparity maps: (Top) with no correction, (Left) with a linear correction model and (Right) with DHW

Optical flow estimation can also benefit from dy-

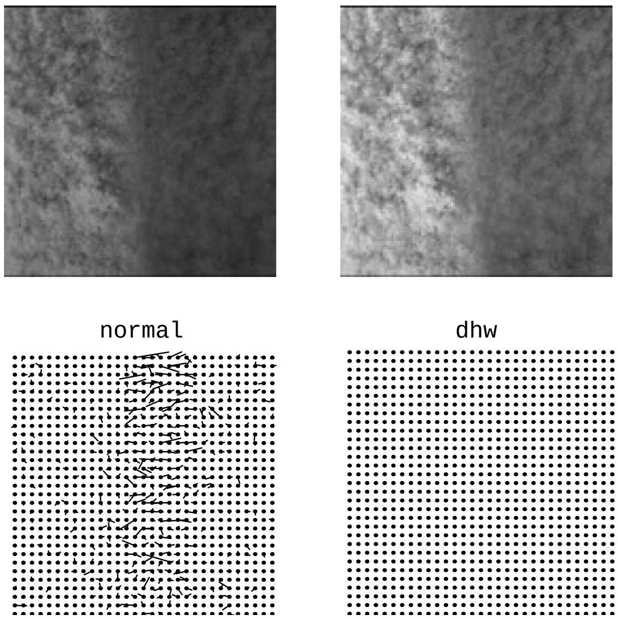

Figure 7: Optical flow for "ceiling" images (top). Flow obtained for original image pair (left). Flow obtained after using DHW (right).

namic histogram warping. Figure (7) shows the image of an office ceiling, courtesy of S. Negahdaripour. "After the first image was taken, the camera aperture was increased before taking the second image" [6].

Figure (7) shows that the optical flow obtained using the method of Horn and Schunk [4] is heavily corrupted because the constant brightness assumption is not satisfied. After correcting the images with DHW, the resulting flow (Figure 7, bottom right) is very close to the real flow, which should be zero everywhere.[4]

4. CONCLUSION

Despite the fact that the constant image brightness (CIB) assumption is common to many branches of computer vision, very little work has been directed to testing this hypothesis. Examination of the SRI JISCT stereo database revealed that the common constant image brightness assumption is often erroneous. This deviation is probably due to several factors which, when lumped together, can be represented as an arbitrary non-linear monotonically increasing function. In this case, errors in the constant image brightness assumption can be corrected, or at least reduced, by matching the intensity histograms of the two images.

Conventional histogram specification based on *local* matching cumulative histograms was shown to be problematic since errors propogate and accumulate and must then be anulled by spurious intensity matches. Instead, a dynamic histogram warping is proposed, analogous to dynamic time warping, that works directly on the intensity histograms by expanding or compressing intensity bins. One-to-one and one-to-many mappings are allowed.

DHW is superior to histogram specification and can be used to correct for constant image brightness without requiring intensity derivatives. We demonstrated this by applying a maximum likelihood stereo algorithm to an image pair that originally deviated significantly from the CIB assumption. The experimental results showed that the while the original disparity map contained many errors, a reduction in errors was achieved by first normalizing the images using DHW. A similar experiment was performed for optical flow estimation. The results clearly show that the accuracy of the resulting flow is greatly improved.

[4] It should be noted that for some stereo and optical flow pairs little of no improvement was obtained. Of course, violations of the constant image brightness assumption represent only one of many sources of possible errors in stereo and optical flow algorithms and it is therefore unreasonable to hope that DHW alone would resolve all such problems.

5. REFERENCES

[1] I. J. Cox. A maximum likelihood N-camera stereo algorithm. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 733–739, 1994.

[2] W. A. Gale and K. W. Church. A program for aligning sentences in bilingual corpora. In *29th Annual Meeting of the Association for Computational Linguistics*, pages 177–184, 1991.

[3] R. C. Gonzalez and R. E. Woods. *Digital Image Processing.* Addison-Wesley, 1993.

[4] K. Horn and B. Schunck. Determining optical flow. *Artificial intelligence*, 17:185–203, 1981.

[5] J. B. Kruskal and M. Lieberman. The symmetric time-warping problem: From continuous to discrete. In D. Sankoff and J. B. Kruskal, editors, *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison.* Addison Wesley, 1983.

[6] Shahriar Negahdaripour and Chih-Ho Yu. A generalized brightness change model for computing optical flow. In *Proc. 4th Int. Conference on Computer Vision*, pages 2–11, Berlin, 1993.

[7] X-D Yang, Q. Xiao, and H. Raafat. Direct mapping between histograms: An improved interactive image enhancement method. In *IEEE Int. Conf. on Systems, Man and Cybernetics*, pages 243–247, 1991.

[8] Y. J. Zhang. Improving the accuracy of direct histogram specification. *Electronics Letters*, 28(3):213–214, 1992.