An Evaluation of the Constant Image Brightness Assumption and its Correction by Dynamic Histogram Warping

Ingemar J. Cox, Sébastien Roy*
NEC Research Institute
4 Independence Way
Princeton, NJ 08540, U.S.A.
ingemar|sebastien@research.nj.nec.com
tel:(609) 951-2722 , fax:(609) 951-2482

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 (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 the many of these algorithms are very sensitive to deviations from CIB. While it is commonly believed that the image brightness assumption is false, it is usually assumed that this deviation can be modelled by a simple global spatially-invariant additive constant and/or a global spatially-invariant scaling of the image intensities (contrast).

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 and linear models 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 might be represented by a non-linear monotonically increasing function of 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 constant image brightness assumption and correct for it. The dynamic histogram warping (DHW) is performed via dynamic programming and is closely related to histogram specification. However, while histogram specification produces good matches, the *local* matching of cumulative histograms introduces artifacts (spikes in the matched histograms) because matching errors propogate and accumulate and must periodically be corrected. This problem does not occur with dynamic histogram warping, in which a global minimum is found.

Experimental results show that image histograms are closely matched after DHW. A further reason for this is that while histogram specification only modifies one histogram DHW can modify both histograms simulatenously. This is especially useful when expansion of an intensity bin of one histogram is not possible but a corresponding compression of the other histogram is. DHW is also capable of removing simple constant additive and multiplicative biases without derivative operations, thereby avoiding amplification of high frequency noise.

EDICS Category : IP 1.4 (Restoration)

*Visiting from Université de Montréal, Département d'informatique et de recherche opérationnelle C.P. 6128, Succ. Centre-Ville, Montréal, Québec, Canada, H3C 3J7

1 Introduction

The constant image brightness 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 [14, 9, 5]. However, surprisingly little experimental work has been performed to support this hypothesis, despite the fact the many of these algorithms are very sensitive to deviations from this assumption.

While it is widely believed that the image brightness assumption is seldom true, it is usually assumed that this deviation can be modelled by a simple global spatially-invariant additive constant, i.e. $I_A = I_B + \alpha$ where I_A and I_B are the intensities of corresponding points in a pair of images. In this case, image *contrast* is conserved and the DC bias is usually removed by applying a first derivative operation to both images. Any derivative operation does, of course, amplify high frequency noise and this can pose a problem for noisy images.

For pixel-based stereo, Gennert [7] provided a detailed model of the intensity relationship between corresponding pixels, showing that corresponding intensities in the left and right images differ by a spatially varying multiplicative factor due to surface orientation and reflectance models. Later work by Cox *et al* [4, 3], however, suggested that this relationship was over shadowed by global changes in illumination conditions and differences in camera responses that were probably the principal source of errors to the constant image brightness assumption. These changes in illumination and or camera responses were modeled by constant multiplicative and additive factors, i.e. $I_A = \beta I_B + \alpha$, that were automatically estimated by a simple analysis of the image histograms. Negahdaripour and Yu [12] have recently proposed a "generalized brightness change" model for optical flow in which the additive and multiplicative terms are spatially varying. This is a very general model for optical flow in which both the x, y flow field and the additive and multiplicative intensity relationships must be estimated at every pixel. The work described here is simpler but restricted to the case where a global, spatially invariant, non-linear, monotonically increasing relationship exists between the intensities of the two images.

This paper first examines the images contained in the SRI JISCT stereo database [2]. Section (2) reveals that the constant image brightness assumption is often false. Moreover, neither a DC bias nor a linear model adequately represents the observed relationships. A comprehensive physical model of the observed deviations is difficult to develop. In fact, it is unlikely that a single model will explain all such relationships, particularly if much of the deviations are attributable to highly nonlinear automatic gain controls common to most manufacturers' cameras.

Clearly, when the CIB assumption is valid, the intensity histograms for a pair of stereo or motion sequence images should be identical, ignoring noise and occlusion affects. However, if corresponding intensities are related by an unknown non-linear monotonically increasing relationship then the intensity histograms will suffer corresponding distortions. We propose to estimate and correct for the unknown non-linear distortion by searching for an optimum non-linear warping of one histogram to the other, that minimizes a cost function defined in Section (3). Such a warping should consist entirely of expansion/contraction of the intensity levels in a manner analagous to the dynamic time warping of speech waveforms [11]. These algorithms are related to the dynamic programming algorithms for stereo [1], but differ by replacing deletions/occlusions with expansion/contractions that allow *non-unique* matches. Section (3) develops what we have called the *dynamic histogram warping.*

The work described here is very closely related to work in histogram specification [8]. Traditionally, this has been a two step process in which the two histograms are first equalized. The final mapping is computed by mapping intensities in the first histogram to their equalized value and then inverse mapping from the equalized value to the corresponding intensity value of the second histogram. Yang *et al* [15] point out that because of quantization errors, this two-step algorithm can produce contouring artifacts. Instead, they propose a direct method that matches the pair of histograms such that each intensity level i is mapped to a corresponding intensity j that minimizes $\left|H_i^A - H_j^B\right|$, where H_i is the cumulative histogram for the fist i intensities. This approach significantly reduces artifacts due to contouring. Zhang [16] showed that the direct method (called SML) can produce poor results because each source intensity is independently mapped to a destination intensity. Zhang suggested modifying the cost function to $\left|H_{f(j)}^A - H_j^B\right|$, where $f(j)$ is a monotonic mapping function, to take advantage of the fact that in general the destination histogram contains less intensities than the source histogram. This method for histogram specification (called GML) produces good matches and the computation is simple. However, the local search for matches introduces artifacts (spikes in the matched histograms) because matching errors accumulate. The simple example of Figure (1) illustrates

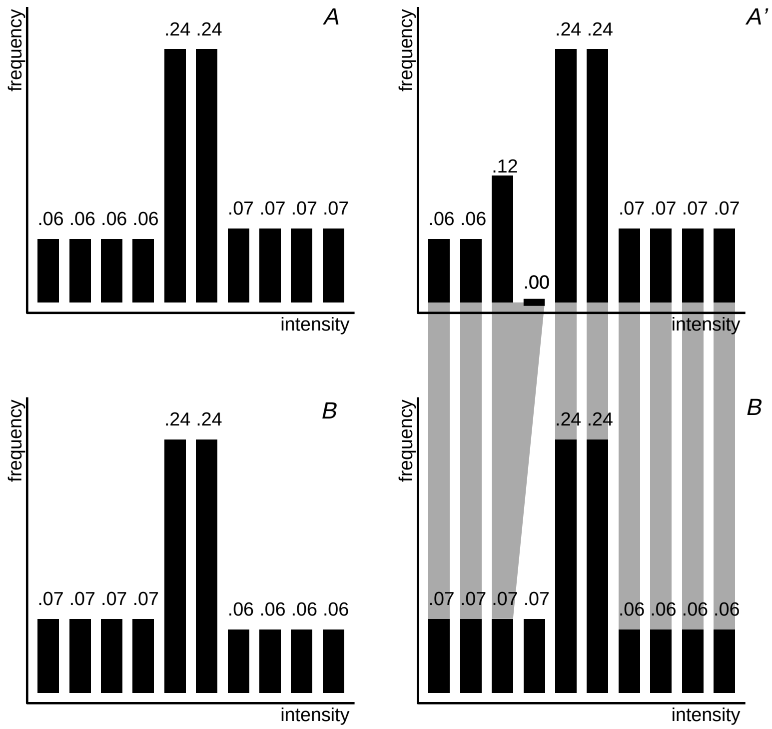

Figure 1: Incorrect histogram specification. A) Original histogram. B) Specified histogram. A') Resulting histogram using SML histogram specification as in [16]. Using GML also gives erroneous results for this case.

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. A further distinction betwwen the two procedures is that while histogram specification only alters one histogram, DHW allows both histograms to be simultaneously modified. The benefit of this is obvious when an intensity bin in one histogram needs to be expanded to many bins in the other. In practice, this is not possible. However, an expansion of one histogram implies a corresponding compression of the other histogram, which is easily accomplished when both histograms are allowed to be modified.

Section (4) describes experimental results of applying the DHW to image histograms. Very close matching is achieved. These results are compared with those achieved through histogram specification and it is clearly shown that DHW produces closer matches and fewer large errors that appear as spikes.

2 The Constant Image Brightness Assumption and the JISCT Database

The SRI JISCT database is a collection of 49 stereo pairs from five sites: JPL, INRIA, SRI, CMU and Teleos. To determine the accuracy of the CIB assumption, we examined corresponding intensity histograms for each of the stereo pairs, since if the CIB assumption is valid, their intensity histograms should be almost identical. Figures (2-5) show that for 15 cases, the CIB assumption was clearly not valid. Vertical lines have been drawn between corresponding points in each pair of histograms: the longer the line, the bigger the difference between the histograms at that pont. The differences between corresponding histograms reveals that a DC bias, i.e. $I_A = I_B + \alpha$, is not an accurate model of the intensity relationship. This is clear from the histograms of "J2" in Figure (4) in which the dark intensity portion of the histogram is well matched but progressively greater deviations are apparent for brighter intensities. Also, the histograms of the the "ARROYO" pair of Figure (2) are clearly structurally different.

To determine the validity of the constant additive and multiplicative model, i.e. $I_A = \beta I_B + \alpha$, the derivative of each image was first taken to remove the DC bias. Then, the histograms of the logs of the magnitude of the image derivatives were computed. If this were an accurate model, then corresponding histograms of the log of absolute derivatives should be shifted versions of one another. Figures (6-9) clearly show that this is not the case. In particular, several of the figures appear to be scaled versions of one another, see Figure (9) for example, indicating that some form of power law relationship may be present.

This analysis shows that not only is the constant image brightness assumption often invalid, but also that the simple constant or linear models for the deviation do not adequately represent the imaging process. The analysis suggests that an alternative model for the relationship between the two sets of intensities might be a non-linear model of the form $I_A = \beta I_B^{\gamma} + \alpha$.

3 Dynamic Histogram Warping (DHW)

There are a number of possible reasons why a pair of images might deviate from the CIB assumption for some scenes, 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

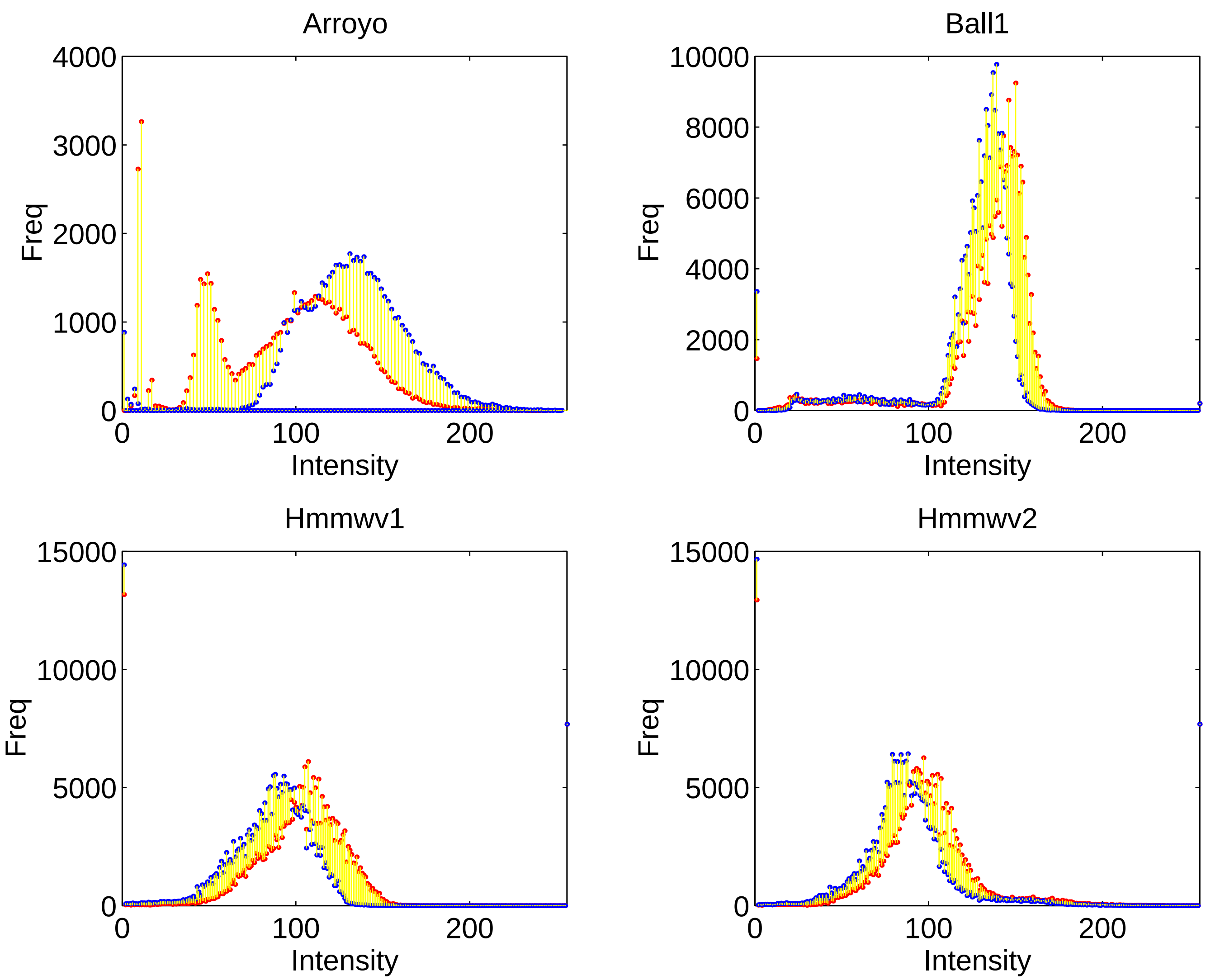

Figure 2: Intensity histograms for JISCT images pairs. Vertical lines have been drawn between corresponding points in each pair of histograms: the longer the line, the bigger the difference between the histograms at that point.

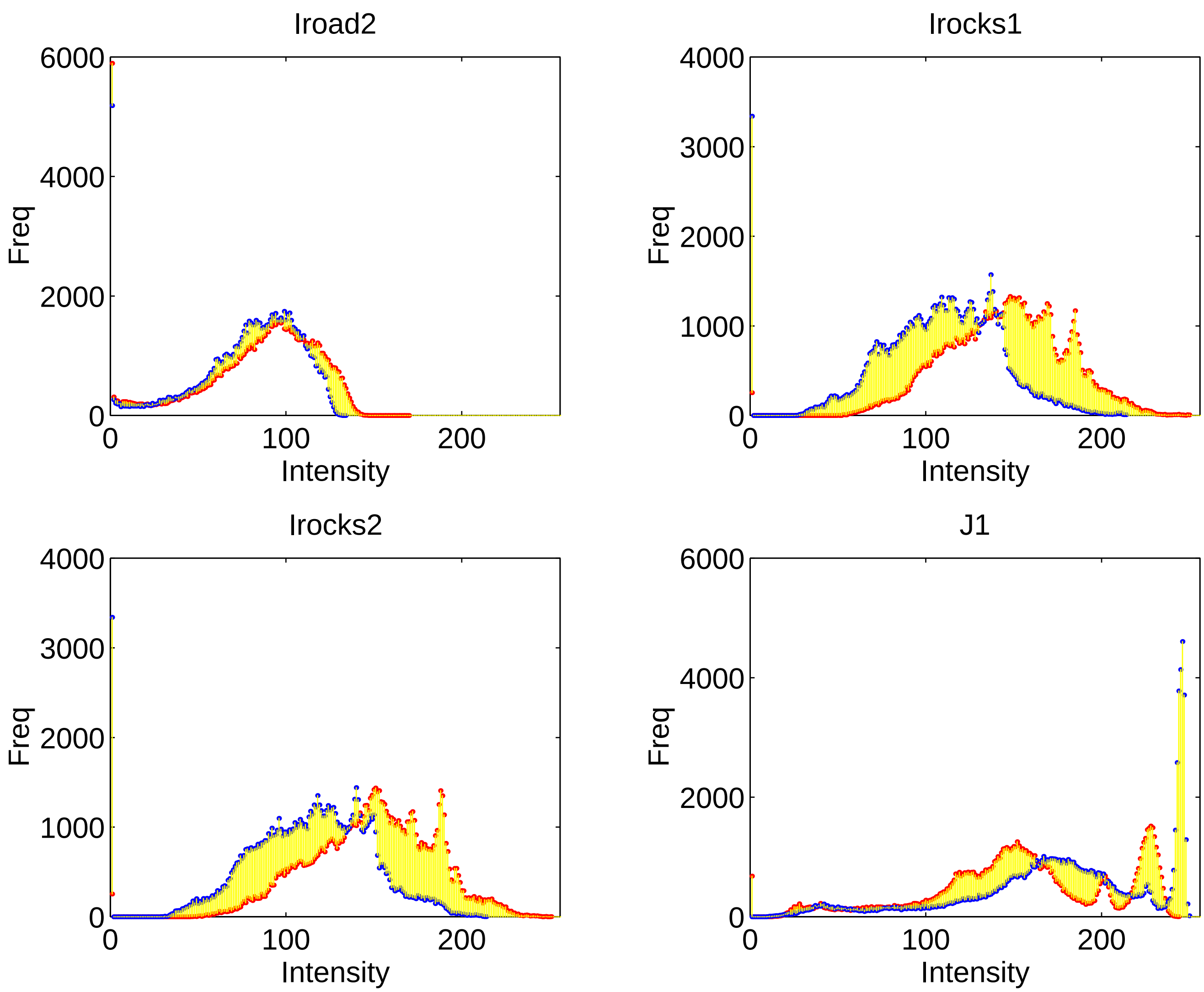

Figure 3: Intensity histograms for JISCT images pairs. Vertical lines have been drawn between corresponding points in each pair of histograms: the longer the line, the bigger the difference between the histograms at that point.

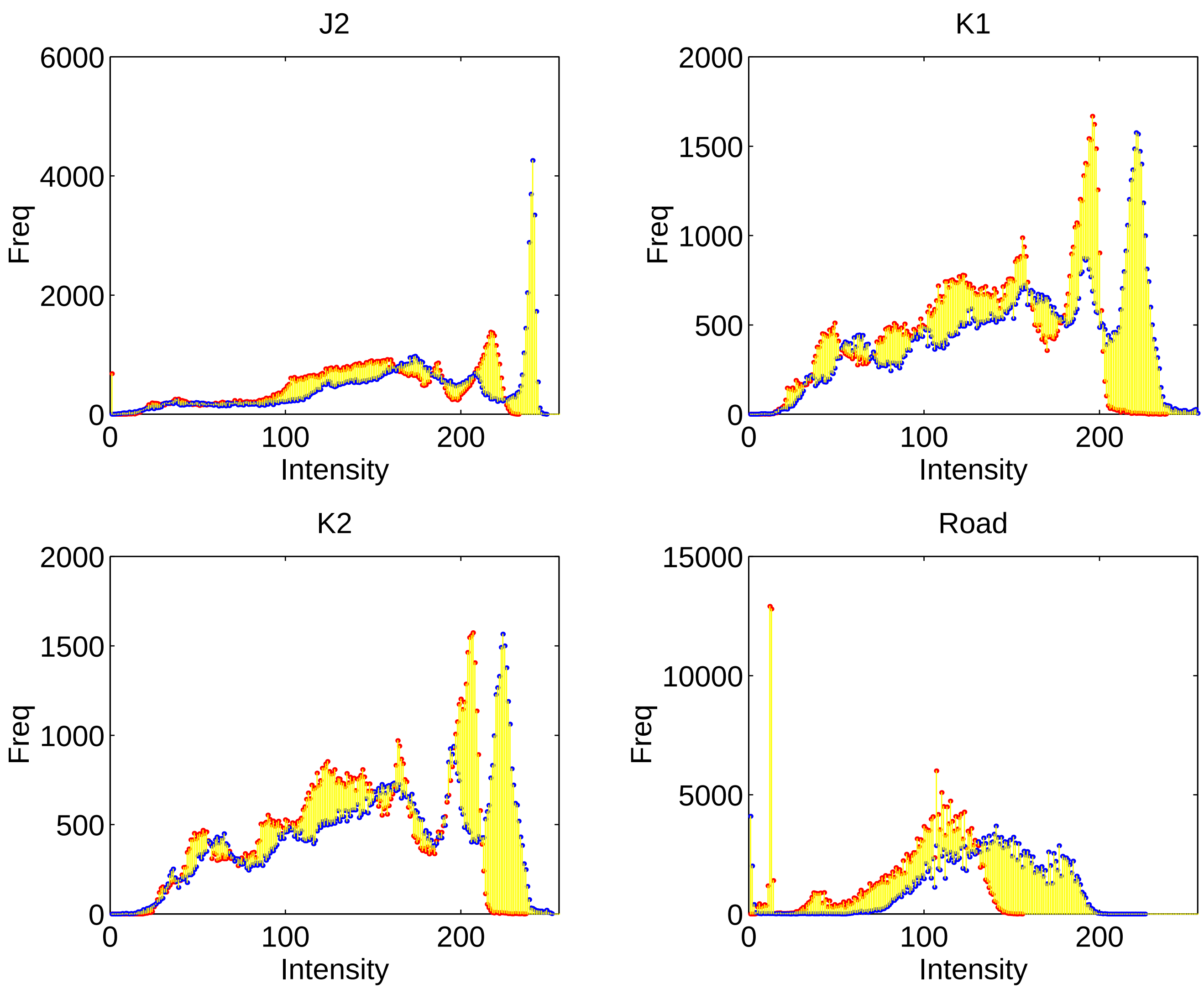

Figure 4: Intensity histograms for JISCT images pairs. Vertical lines have been drawn between corresponding points in each pair of histograms: the longer the line, the bigger the difference between the histograms at that point.

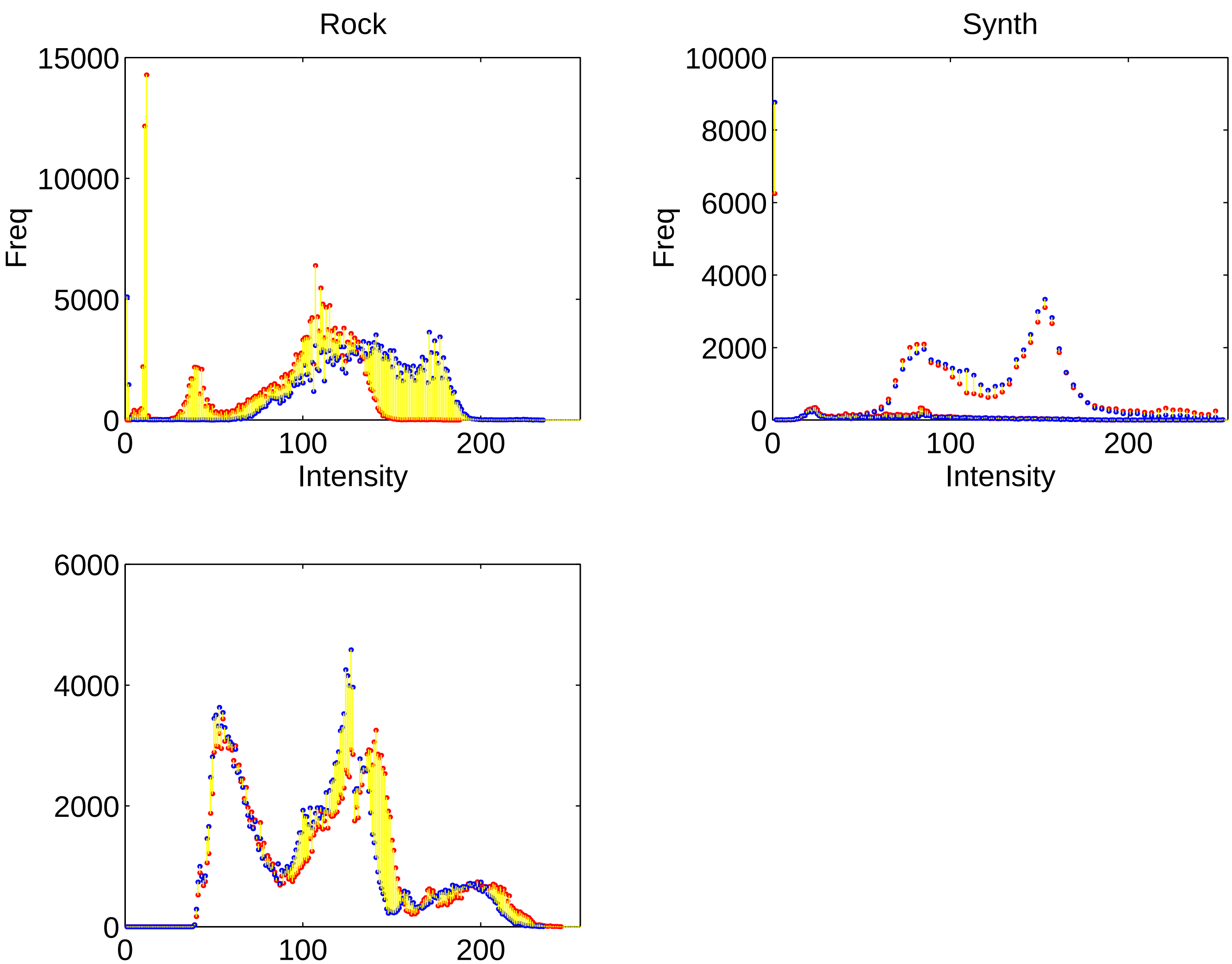

Figure 5: Intensity histograms for JISCT images pairs. Vertical lines have been drawn between corresponding points in each pair of histograms: the longer the line, the bigger the difference between the histograms at that point.

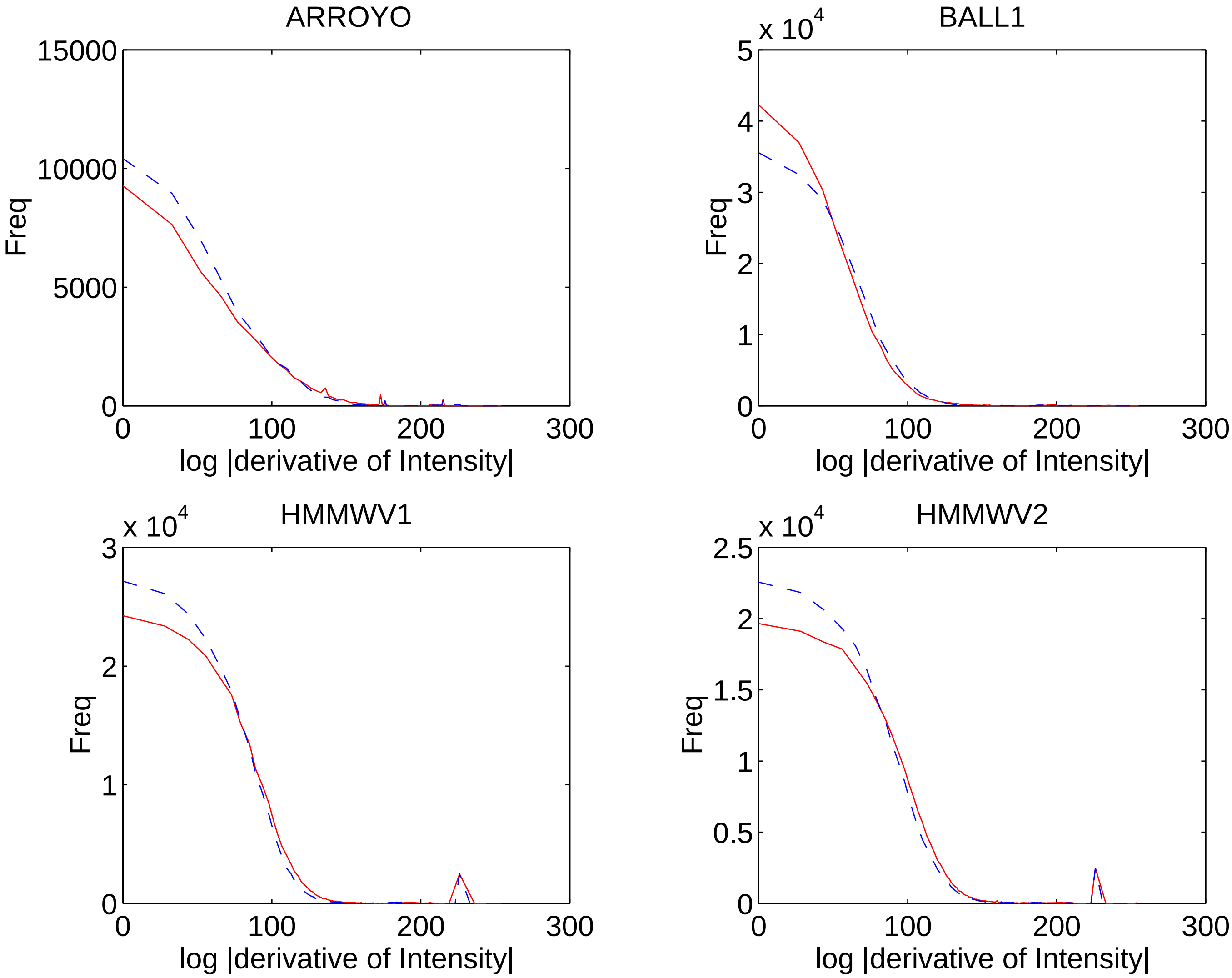

Figure 6: Normalized log derivative intensity histograms for sample JISCT images pairs.

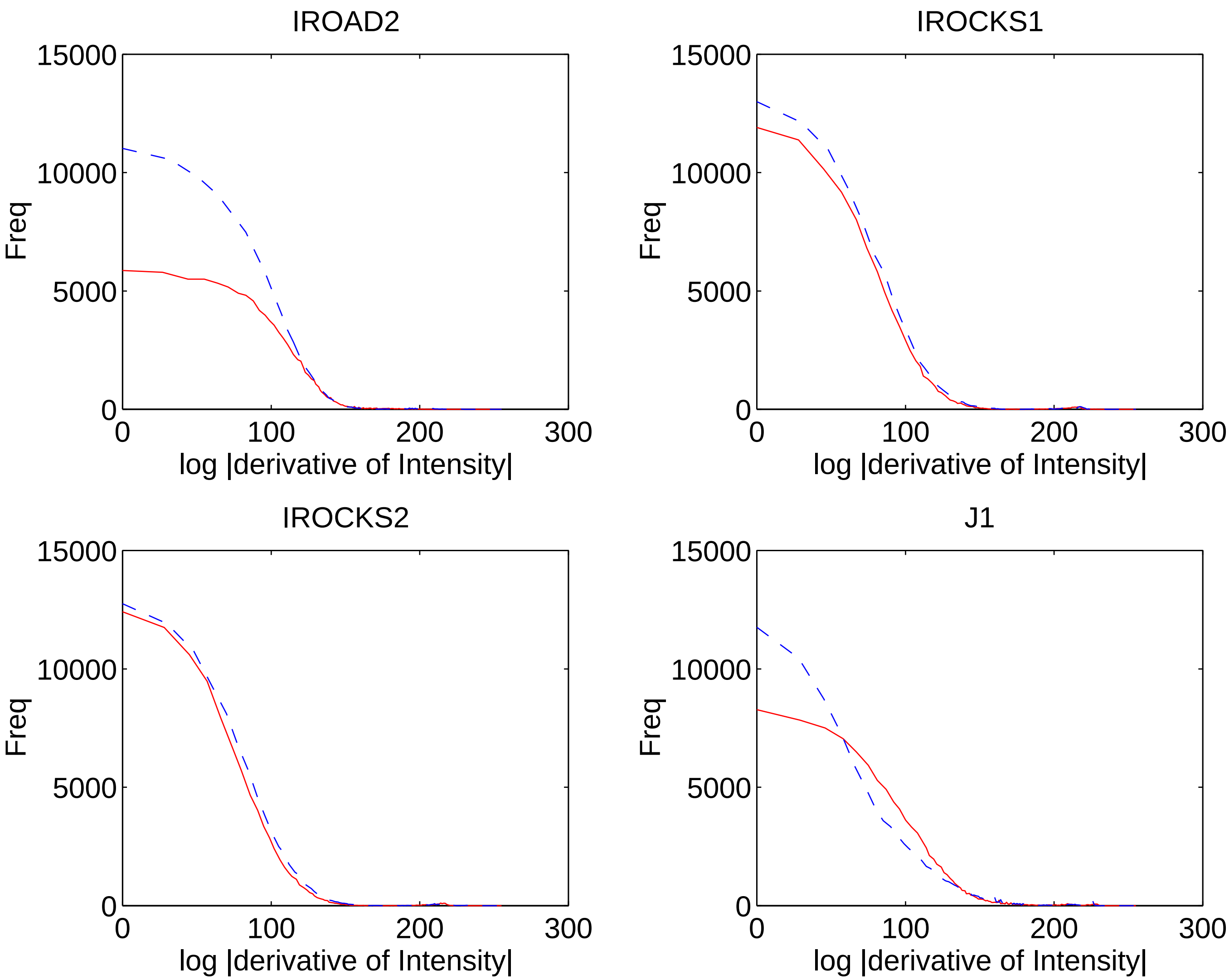

Figure 7: Normalized log derivative intensity histograms for sample JISCT images pairs.

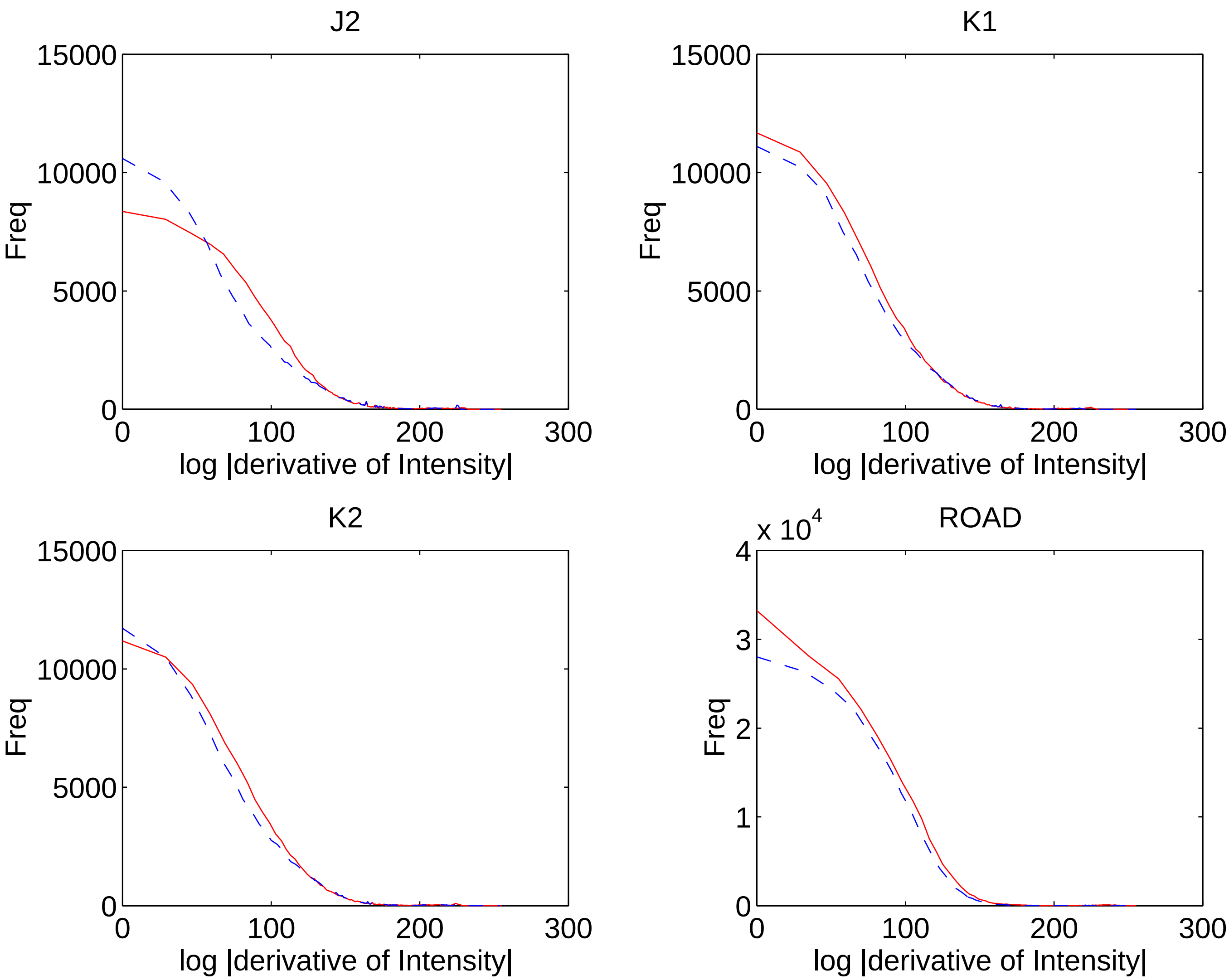

Figure 8: Normalized log derivative intensity histograms for sample JISCT images pairs.

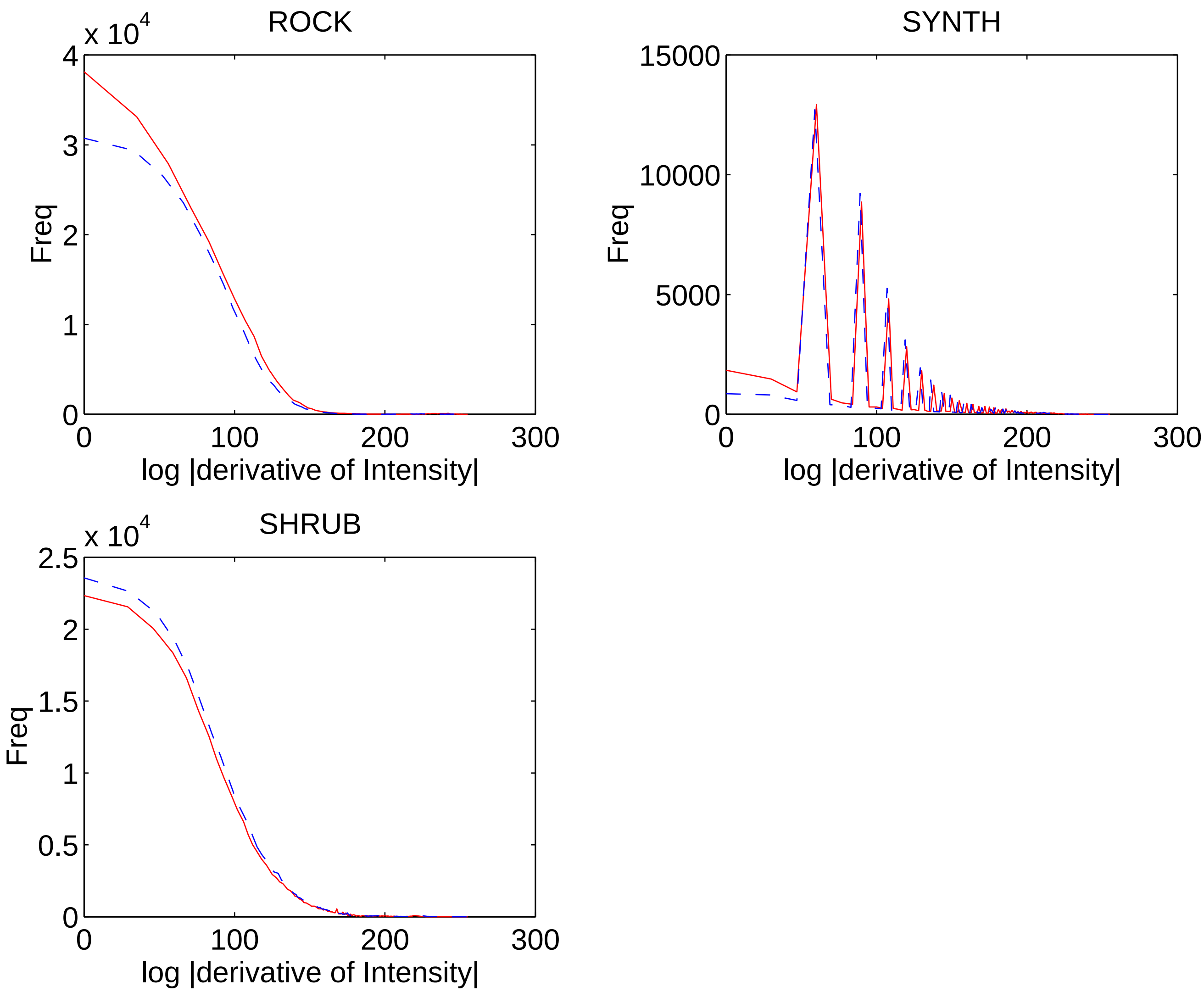

Figure 9: Normalized log derivative intensity histograms for sample JISCT images pairs.

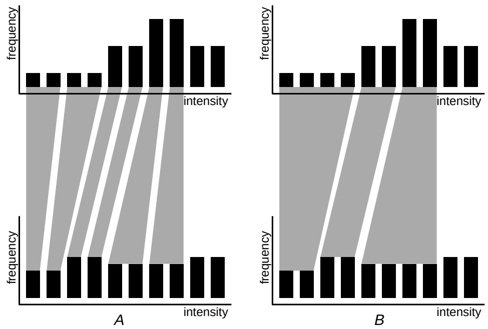

Figure 10: Histogram matching. Legal matches (A) always join only one intensity to one or more others. Illegal matches (B) join many intensities to many others.

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. Such a comparison and correction is strongly related to work in sequence comparison [13] and especially to dynamic time warping (DTW), commonly used in speech recognition to minimize variations in the rate of speech between speakers [11]. In dynamic time warping, two speech signals are compressed and/or expanded to best match one another. Signal samples can be matched one-to-one, one-to-many (expansion) or many-to-one (contraction), as illustrated in Figure (10a). However, the many-to-many mappings of Figure (10b) are illegal.

Because of quantization error, for example, 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 decided to follow DTW and not allow many-to-many mappings.

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 programming. However, since the maximum size of a compression or expansion is always finite[1], then such a cost function can be accomodated [6]. In general, the cost of a k-to-l mapping is

$$d_{k,l}(m,n) = \left|\sum_{i=0}^{k-1} h_{m-i}^A - \sum_{j=0}^{l-1} h_{n-j}^B\right| = \left|(H_m^A - H_{m-k}^A) - (H_n^B - H_{n-l}^B)\right|$$

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

$$\begin{aligned} &D(0,0) = 0 \\ &D(i,j) = \infty \qquad (i \leq 0, j \leq 0, (i,j) \neq (0,0)) \\ &D(m,n) = Min\,[D(m-k, n-l) + d_{k,l}(m,n)] \quad \begin{array}{c} 1 \leq k \leq M \\ 1 \leq l \leq N \\ (k-1)(l-1) = 0 \end{array} \end{aligned}$$

where M and N represent the maximum allowable compression of the respective histograms and the constraint that $(k-1)(l-1) = 0$ prevents many-to-many mappings. The cost function can be efficiently minimized via dynamic programming.

Traditional histogram specification assumes one histogram has a *fixed* reference. DHW is also capable of such constructions but is also more flexible since it is possible to simultaneously warp both histograms, replacing expansions of one histogram by corresponging compressions of the other. The experimental results reported next simultaneously warp both histograms.

4 Experimental results using DHW

The dynamic histogram warping algorithm was applied to the images whose original histograms were shown in Figures (2-5). Figures (11-14) shows the resulting histograms after image matching. It is clear that very close matching has been achieved.

In comparison, Figures (15-18) show the results of applying a conventional histogram specification (GML) algorithm to match intensity histograms. Although reasonably good matching is achieved, spurious matches, in the form of spikes, are clearly visible.

The difference between resulting histograms, measured as a sum of squared differences ($e = \sqrt{\sum_i (h_i'^A - h_i'^B)^2}$), is shown in Figure (19) for three methods applied to the 15 test image

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

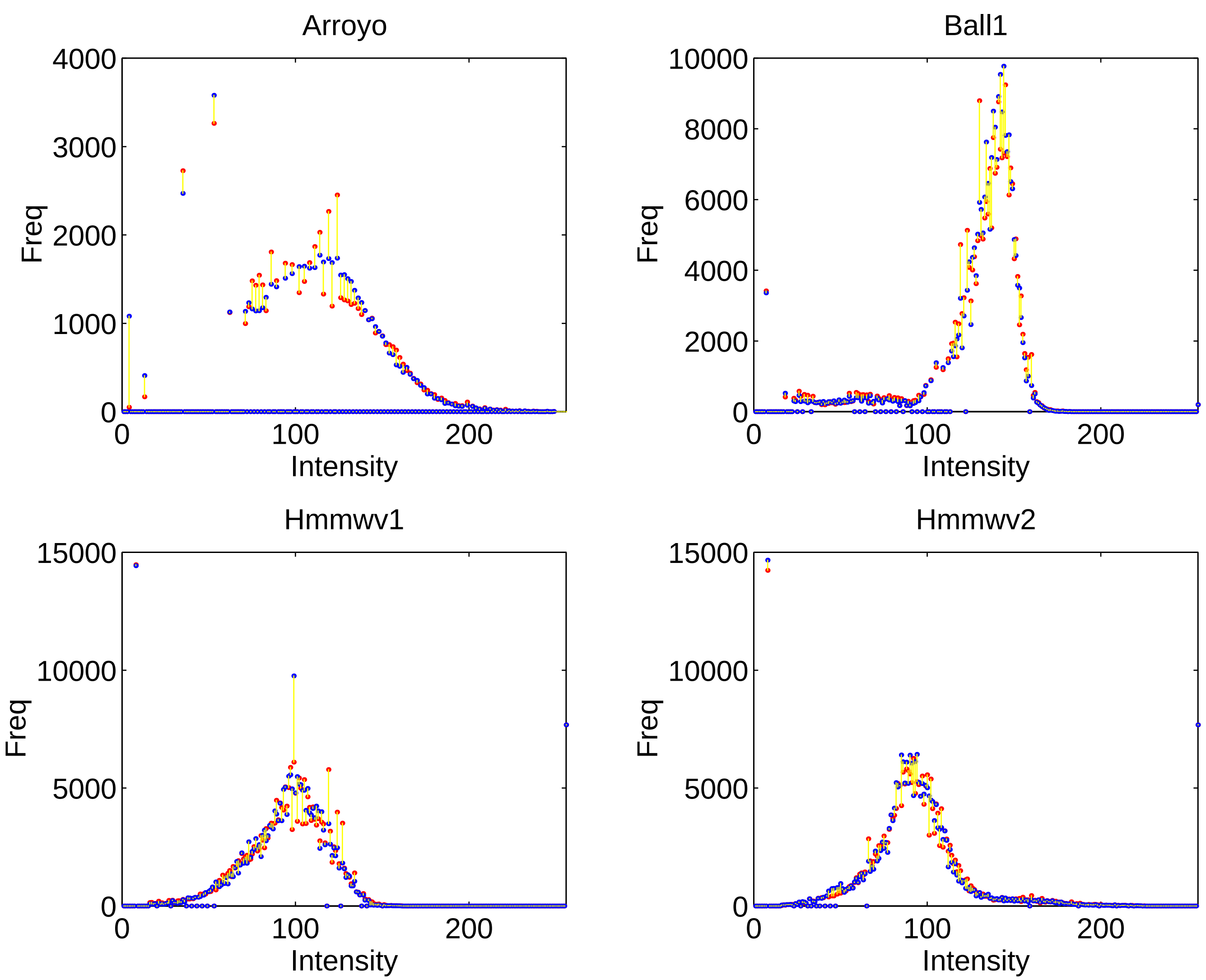

Figure 11: Intensity histograms for JISCT images pairs after dynamic histogram warping.

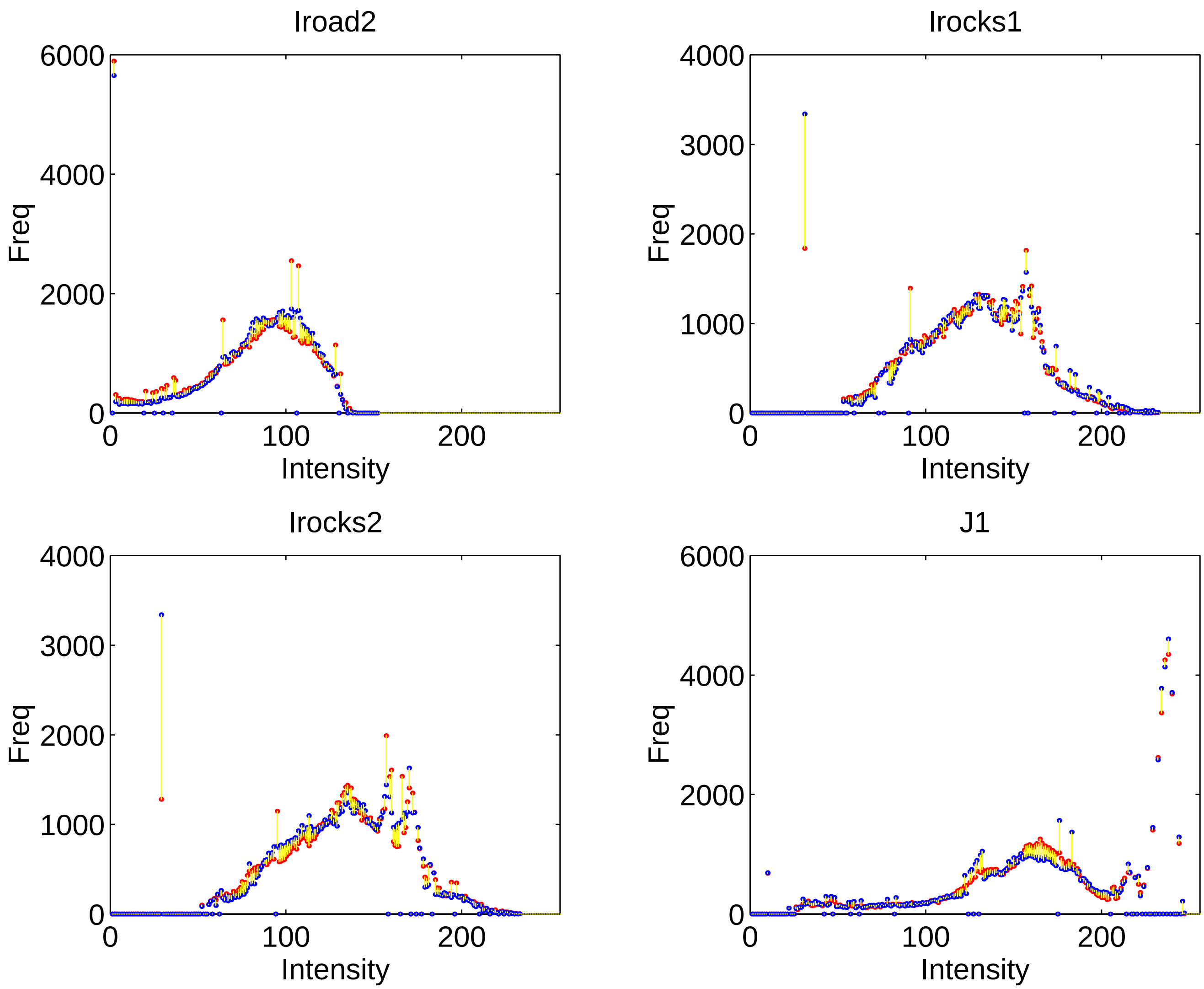

Figure 12: Intensity histograms for JISCT images pairs after dynamic histogram warping.

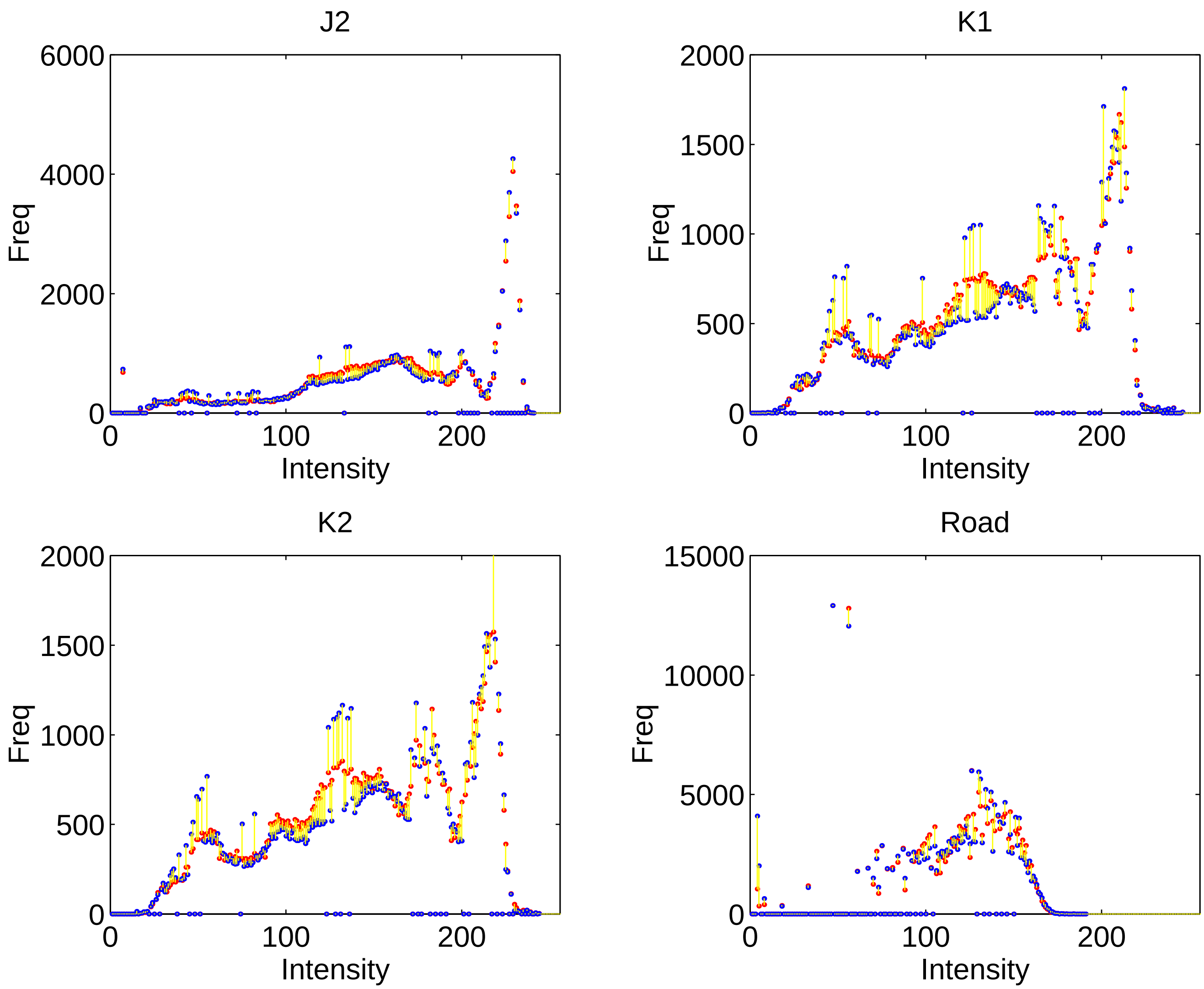

Figure 13: Intensity histograms for JISCT images pairs after dynamic histogram warping.

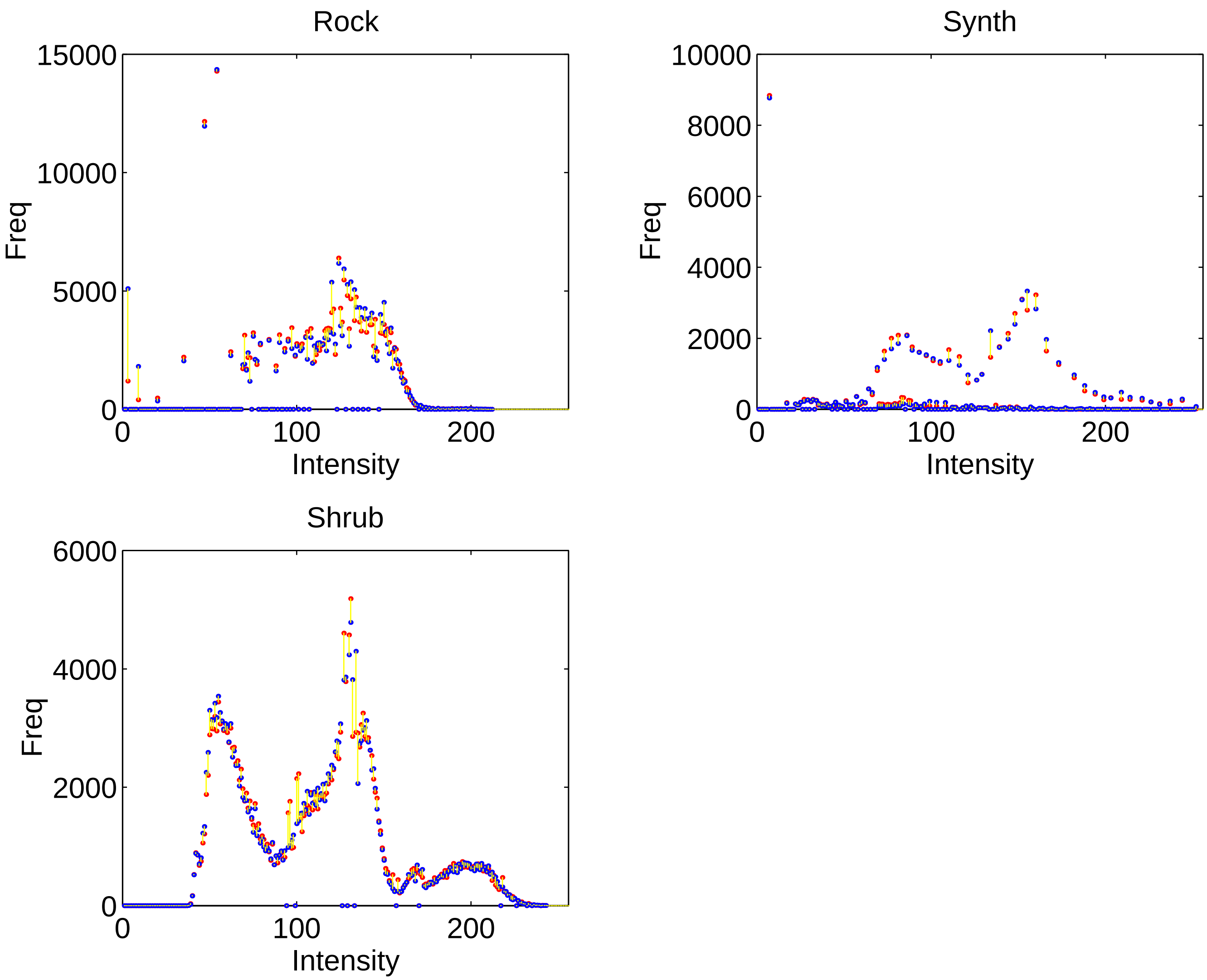

Figure 14: Intensity histograms for JISCT images pairs after dynamic histogram warping.

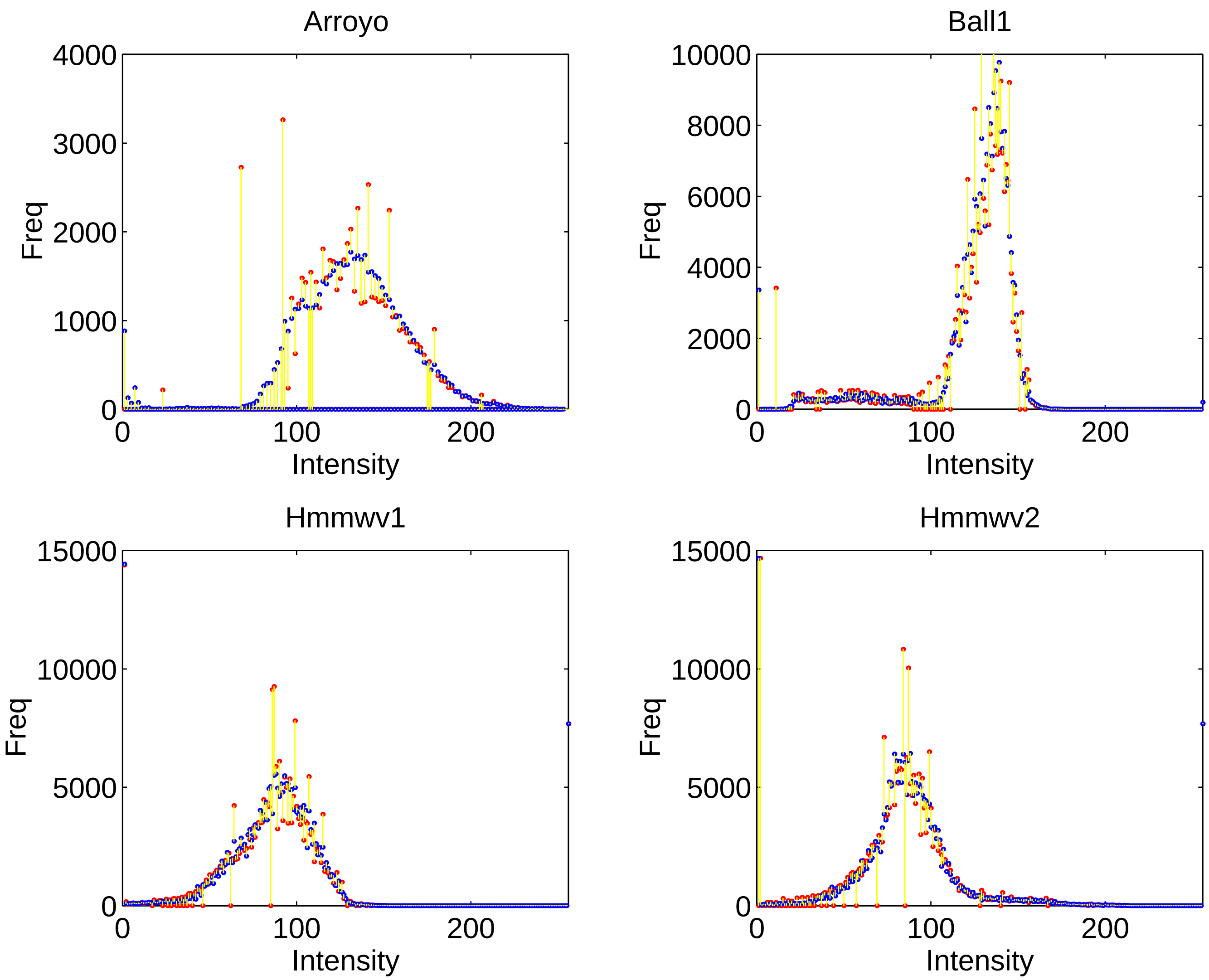

Figure 15: Intensity histograms for JISCT images pairs after histogram specification.

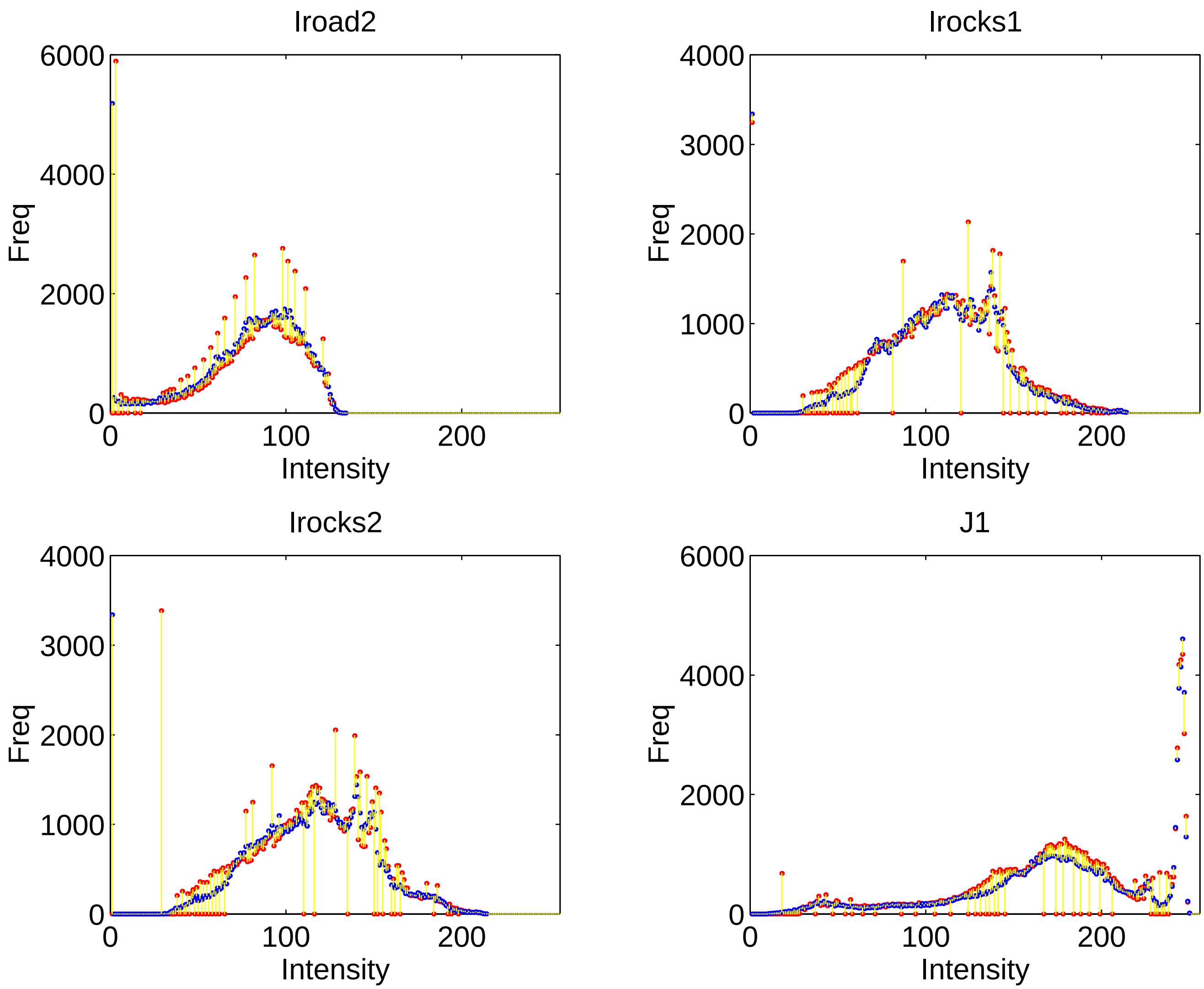

Figure 16: Intensity histograms for JISCT images pairs after histogram specification.

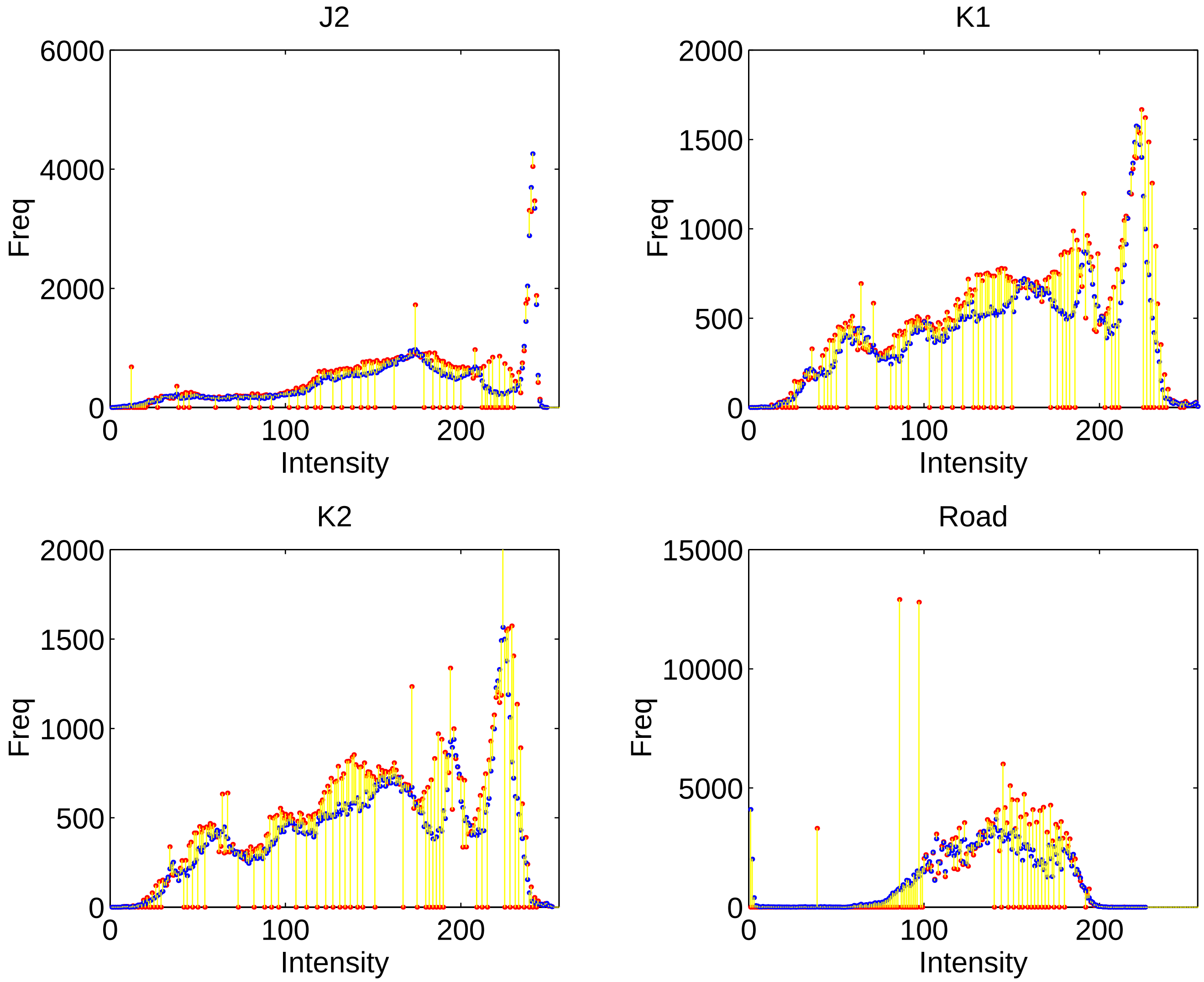

Figure 17: Intensity histograms for JISCT images pairs after histogram specification.

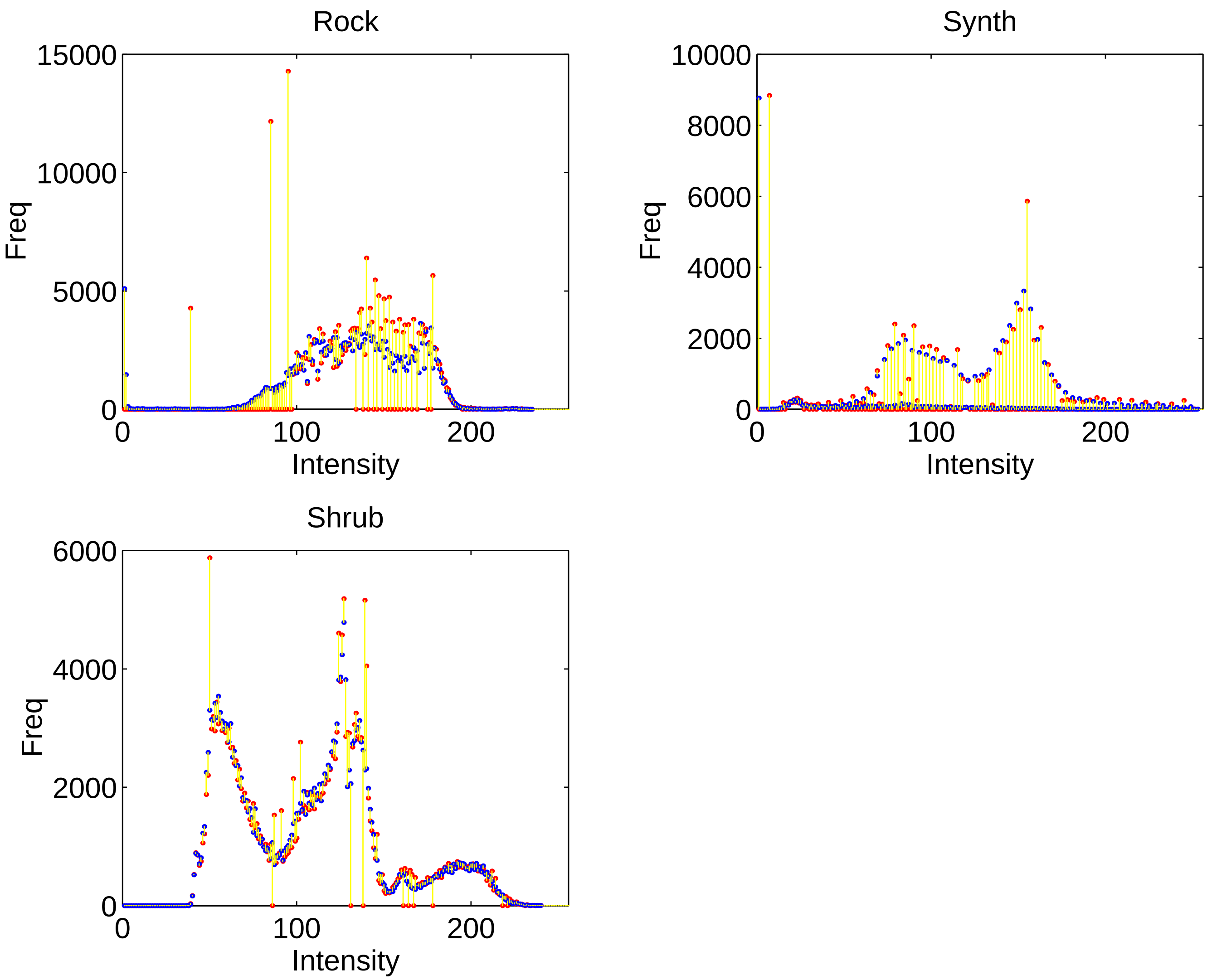

Figure 18: Intensity histograms for JISCT images pairs after histogram specification.

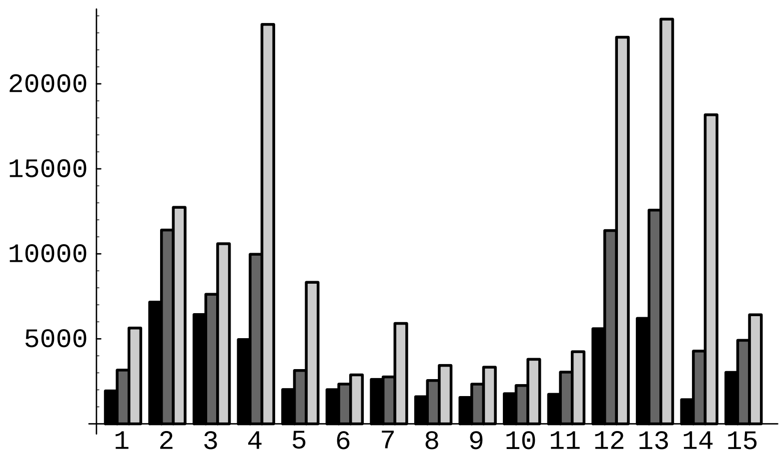

Figure 19: Sum of square differences. Black: Dynamic Histogram Warping. Dark gray: regular normalisation (GML). Light gray: regular histogram specification (GML)

pairs. The dynamic histogram warping (in black) always yield a smaller error then regular normalisation or specification.

As an example of the type of correction achieved, Figure (20) shows the original pair of images, "IROAD2", from the JISCT database. We applied a stereo algorithm [4, 3] to this pair (1) with no correction, (2) correction using a linear model of the intensity relationship, (3) correction using regular histogram specification and (4) correction using dynamic histograms warping. The corresponding disparity maps are shown in Figure (21).

With no histogram correction, the disparity map is extremely poor. While the linear model substantially improves the disparity output, some artifacts are present, mostly on the upper left portion of the road. Both histogram specification and DHW remove this artifact but histogram specification introduces some additional error in the left foreground of the road.

Optical flow estimation can also benefit from dynamic histogram warping. Figure (22) 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" [12].

The optical flow obtained using the method of Horn & Schunk [10] is shown in Figure (23a) and is heavily corrupted because the constant brightness assumption is not satisfied. After correcting the images with DHW, the resulting flow shown in Figure (23b) is very close to the real flow, which should be zero everywhere.

Figure 20: The "IROAD2" stereo pair, included in the SRI JISCT database.

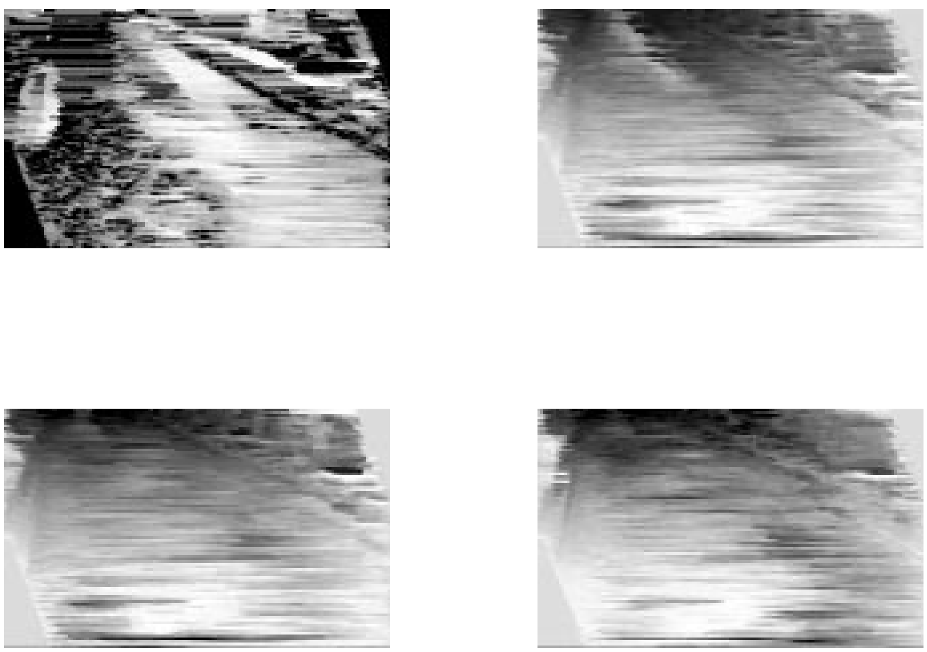

Figure 21: Disparity maps for the "IROAD2" stereogram. Upper left: No correction. Upper right: Linear correction model. Lower left: regular histogram specification. Lower right: Dynamic Histogram Warping.

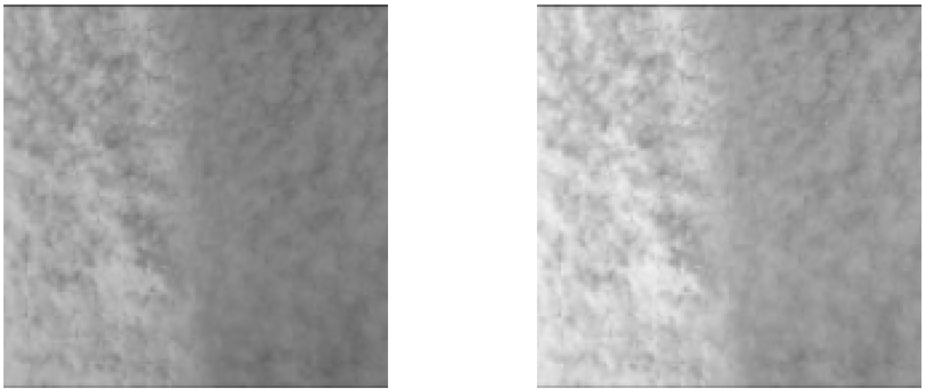

Figure 22: Office ceiling image pair. Courtesy of S. Negahdaripour

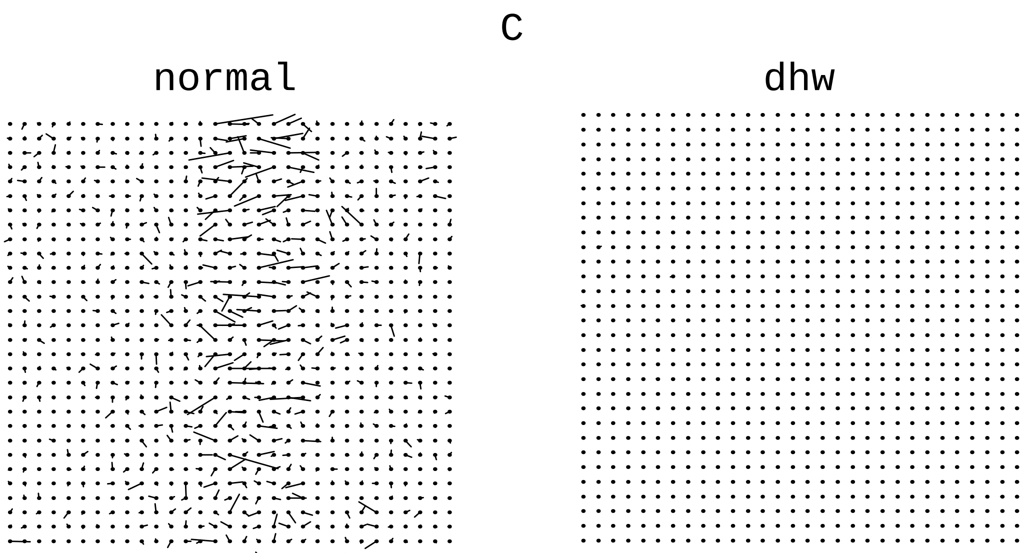

Figure 23: Optical flow. A) Flow obtained for original image pair. B) Flow obtained after using DHW

5 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. Forty nine image pairs from the SRI JISCT stereo database were examined and it was empirically demonstrated that the CIB assumption is often erroneous. Inspection of corresponding histogram pairs revealed that the common additive (DC bias) and linear ($I_A = \beta I_B + \alpha$) models are not good models of the intensity relationship between the two images.

The deviation from constant image brightness is probably due to several factors including variations in illumination, camera responses and non-linear automatic gain controls. If these factors are 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.

Conventional histogram specification based on *local* matching corresponding 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.

Dynamic histogram warping was applied to 14 image pairs from the SRI JISCT database that had previously been identified as not meeting the CIB assumption. An examination of the corrected histograms indicated very close matchings that were superior to those achievable by conventional histogram specification.

DHW was designed as a front-end preprocessing stage to computer vision algorithms that assume constant image brightness. 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.

The constant image brightness assumption is central to many computer vision algorithms and the dynamic histogram warping algorithm developed here provides a powerful method for measuring the validity of the CIB assumption and correcting for deviations. Moreover, because DHW does not require image derivatives, it does not amplify high frequency noise. As a result, one may be able to apply existing stereo, optical flow and color indexing algorithms to noisier imagery than was previously possible. It is hoped that the dynamic histogram warping will become a standard preprocessing stage for algorithms that assume constant image brightness.

Acknowledgements

The authors thank J. B. Kruskal of AT&T Bell Labs for very useful discussions. Also, thanks to Indranil Chakravarty and Ajit Singh of Siemens Corporate Research.

References

[1] H. H. Baker and T. O. Binford. Depth from edge and intensity based stereo. In *Int. Joint Conf. on Artificial Intelligence*, pages 631–636, 1981.

[2] R. C. Bolles, H. H. Baker, and M. J. Hannah. The JISCT stereo evaluation. In *Proc. of DARPA Image Understanding Workshop*, pages 263–274, 1993.

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

[4] I. J. Cox, S. Hingorani, B. M. Maggs, and S. B. Rao. Stereo without disparity gradient smoothing: a Bayesian sensor fusion solution. In D. Hogg and R. Boyle, editors, *British Machine Vision Conference*, pages 337–346. Springer-Verlag, 1992.

[5] B. V. Funt and G. D. Finlayson. Color constant color indexing. *??*, 1994.

[6] 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.

[7] M. A. Gennert. *A Computational Framework for Understanding Problems in Stereo Vision.* PhD thesis, MIT, 1987.

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

[9] G. Healey and D. Slater. Global color constancy: recognition of objects by use of illumination-invariant properties of color distributions. *J. Optical Society of America A*, 1(11):3003–3010, 1994.

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

[11] 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.

[12] 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.

[13] D. Sankoff and J. B. Kruskal. *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison.* Addison Wesley, 1983.

[14] M. J. Swain and D. H. Ballard. Color indexing. volume 7, pages 11–32, 1991.

[15] 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.

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