The Quadratic-Chi Histogram Distance Family

Ofir Pele and Michael Werman

School of Computer Science
The Hebrew University of Jerusalem
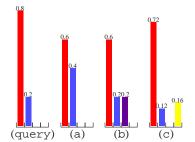{ofirpele,werman}@cs.huji.ac.il

Abstract. We present a new histogram distance family, the Quadratic-Chi (QC). QC members are Quadratic-Form distances with a cross-bin χ^2-like normalization. The cross-bin χ^2-like normalization reduces the effect of large bins having undo influence. Normalization was shown to be helpful in many cases, where the χ^2 histogram distance outperformed the L_2 norm. However, χ^2 is sensitive to quantization effects, such as caused by light changes, shape deformations etc. The Quadratic-Form part of QC members takes care of cross-bin relationships (*e.g.* red and orange), alleviating the quantization problem. We present two new cross-bin histogram distance properties: *Similarity-Matrix-Quantization-Invariance* and *Sparseness-Invariance* and show that QC distances have these properties. We also show that experimentally they boost performance. QC distances computation time complexity is linear in the number of non-zero entries in the bin-similarity matrix and histograms and it can easily be parallelized. We present results for image retrieval using the Scale Invariant Feature Transform (SIFT) and color image descriptors. In addition, we present results for shape classification using Shape Context (SC) and Inner Distance Shape Context (IDSC). We show that the new QC members outperform state of the art distances for these tasks, while having a short running time. The experimental results show that both the cross-bin property and the normalization are important.

1 Introduction

It is common practice to use bin-to-bin distances such as the L_1 and L_2 norms for comparing histograms. This practice assumes that the histogram domains are aligned. However this assumption is violated in many cases due to quantization, shape deformation, light changes, etc. Bin-to-bin distances depend on the number of bins. If it is low, the distance is robust, but not discriminative, if it is high, the distance is discriminative, but not robust. Distances that take into account cross-bin relationships (cross-bin distances) can be both robust and discriminative.

There are two kinds of cross-bin distances. The first is the Quadratic-Form distance [1]. Let P and Q be two histograms and A the bin-similarity matrix. The Quadratic-Form distance is defined as:

$$\mathrm{QF}^A(P,Q) = \sqrt{(P-Q)^T A (P-Q)} \qquad (1)$$

distance	(a)-(c) ordered by their distance (in small fonts) to the (query)					
QCN(our)	(a)	0.35	(b)	0.62	(c)	0.86
QCS(our)	(a)	0.31	(b)	0.41	(c)	0.43
QF	(c)	0.20	(a)	0.28	(b)	0.28
EMD	(c)	3.20	(a)	4.00	(b)	4.00
L_1	(c)	0.32	(a)	0.40	(b)	0.40
χ^2	(a)	0.05	(c)	0.09	(b)	0.11

Fig. 1. This figure should be viewed in color, preferably on a computer screen. A toy example showing the behavior of distances that reduce the effect of large bins and the behavior of distances that take cross-bin relationships into account. We show four color histograms, each histogram has four colors: red, blue, purple, and yellow. The Quadratic-Form (QF), the Earth Mover Distance (EMD) and the L_1 norm do not reduce the effect of large bins. Thus, they rank (query) to be more similar to (c) than to (a). χ^2 considers (a) to be more similar, but as it does not take cross-bin relationships into account it fails with (b). Our proposed members of the Quadratic-Chi histogram distance family, QCN and QCS consider (a) to be most similar, (b) the second and (c) the least similar as they take into account cross-bin relationships and reduce the effect of large bins, using an appropriate normalization.

When the bin-similarity matrix A is the inverse of the covariance matrix, the Quadratic-Form distance is called the Mahalanobis distance. If the bin-similarity matrix is positive-definitive, then the Quadratic-Form distance is a metric. In this case the Quadratic-Form distance is the L_2 norm between linear transformations of P and Q. If the bin-similarity matrix is positive-semidefinite, then the Quadratic-Form distance is a semi-metric.

The second type of distance that takes into account cross-bin relationships is the Earth Mover's Distance (EMD). EMD was defined by Rubner et al. [2] as the minimal cost that must be paid to transform one histogram (P) into the other (Q):

$$\text{EMD}^D(P,Q) = (\min_{\{F_{ij}\}} \sum_{i,j} F_{ij} D_{ij})/(\sum_{i,j} F_{ij}) \quad s.t \quad F_{ij} \geq 0$$

$$\sum_j F_{ij} \leq P_i \quad \sum_i F_{ij} \leq Q_j \quad \sum_{i,j} F_{ij} = \min(\sum_i P_i, \sum_j Q_j) \tag{2}$$

where $\{F_{ij}\}$ denotes the flows. Each F_{ij} represents the amount transported from the ith supply to the jth demand. We call D_{ij} the *ground distance* between bin i and bin j. If D_{ij} is a metric, the EMD as defined by Rubner is a metric only for normalized histograms. Recently Pele and Werman [3] suggested $\widehat{\text{EMD}}$:

$$\widehat{\text{EMD}}_\alpha^D(P,Q) = (\min_{\{F_{ij}\}} \sum_{i,j} F_{ij} D_{ij}) + |\sum_i P_i - \sum_j Q_j| \alpha \max_{i,j} D_{ij} \tag{3}$$

$$s.t \ \text{EMD constraints}$$

If D_{ij} is a metric and $\alpha \geq \frac{1}{2}$, $\widehat{\text{EMD}}$ is a metric for all histograms [3]. For normalized histograms $\widehat{\text{EMD}}$ and EMD are equal (*e.g.* Fig. 1).

In many natural histograms the difference between large bins is less important than the difference between small bins and should be reduced. See for example Fig. 1. The Chi-Squared (χ^2) is a histogram distance that takes this into account. It is defined as:

$$\chi^2(P, Q) = \frac{1}{2} \sum_i \frac{(P_i - Q_i)^2}{(P_i + Q_i)} \tag{4}$$

The χ^2 histogram distance comes from the χ^2 test-statistic [4] where it is used to test the fit between a distribution and observed frequencies. In this paper the histograms are not necessarily normalized, and thus not probabilities vectors. χ^2 was successfully used for texture and object categories classification [5,6,7], near duplicate image identification[8], local descriptors matching [9], shape classification [10,11] and boundary detection [12]. The χ^2, like other bin-to-bin distances such as the L_1 and the L_2 norms, is sensitive to quantization effects.

2   Our Contribution

In this paper we present a new cross-bin histogram distance family: Quadratic-Chi (QC). Like the Quadratic-Form, its members take cross-bin relationships into account. Like the χ^2, its members reduce the effect of differences caused by bins with large values. We discuss QC members' properties, including a formalization of a two new cross-bin histogram distance properties: *Similarity-Matrix-Quantization-Invariance* and *Sparseness-Invariance*. We show that all QC members and the EMD have these properties. We also show importance experimentally.

For full histograms QC distances computation time is linear in the number of non-zero entries in the bin-similarity matrix. In this case, QC distances can be implemented with 5 lines of Matlab code (see Algorithm 1). For two sparse histograms (for example bag-of-words histograms) with a total of S non-zeros entries and an average of K non-zeros entries in each row of the similarity matrix, a QC distance computation time complexity is $O(SK)$. See code (C++ and Matlab wrappers) at: http://www.cs.huji.ac.il/~ofirpele/QC/. Finally, QC distances' parallelization is trivial.

We present results for image retrieval on the Corel dataset using the SIFT descriptor [13] and small color images. We also present results for shape classification using Shape Context (SC) [10] and Inner Distance Shape Context (IDSC) [11]. QC members performance is excellent. They outperform state of the art distances including χ^2, QF, L_1, L_2, $\widehat{\text{EMD}}$[14], SIFT$_{\text{DIST}}$[3], EMD-L_1[15], Diffusion[16], Bhattacharyya [17], Kullback-Leibler[18] and Jensen-Shannon[19] while having a short running time. We have found that the normalization is very important. Surprisingly, excellent performance was achieved using a new bin-to-bin distance from the QC family, that has a large normalization factor. Its cross-bin version yielded an additional improvement, outperforming all other distances for SIFT, SC and IDSC.

3   The Quadratic-Chi Histogram Distance Family

3.1   The Quadratic-Chi Histogram Distance Definition

Let P and Q be two non-negative bounded histograms. That is, $P, Q \in [0, U]^N$. Let A be a non-negative symmetric bounded bin-similarity matrix such that each diagonal element is bigger or equal to every other element in its row (this demand is weaker than being a strongly dominant matrix). That is, $A \in [0, U]^N \times [0, U]^N$ and $\forall i, j \ A_{ii} \geq A_{ij}$. Let $0 \leq m < 1$ be the normalization factor. A Quadratic-Chi (QC) histogram distance is defined as:

$$\mathrm{QC}_m^A(P, Q) = \sqrt{\sum_{ij} \left(\frac{(P_i - Q_i)}{\left(\sum_c (P_c + Q_c) A_{ci} \right)^m} \right) \left(\frac{(P_j - Q_j)}{\left(\sum_c (P_c + Q_c) A_{cj} \right)^m} \right) A_{ij}} \quad (5)$$

where we define $\frac{0}{0} = 0$. If A is positive-semidefinite, the argument inside the square root (the sum) is non-negative. If A is not positive-semidefinite we can get non-real (complex) distances. This is true also for the Quadratic-Form (Eq. 1). We prefer not to restrict ourselves to positive-semidefinite matrices. On the other hand, we don't want non-real distances. So, we define a complex distance as zero. In practice, this was never needed, even with non-positive-semidefinite matrices. This is due to the fact that the eigenvectors of the similarity matrices corresponding to negative eigenvalues were very far from smooth, while the difference vector for natural histograms P and Q is usually very smooth, see Fig. 2.

Each addend's denominator inside the square root is zero if and only if the addend's numerator is zero. A $\mathrm{QC}_m^A(P, Q)$ distance is continuous. In particular, if the addend's denominator tends to zero, the whole addend tends to zero. Proofs are in [20].

The Quadratic-Chi distance family generalizes both the Quadratic-Form (QF) and a monotonic transformation of χ^2. That is, $\mathrm{QC}_0^A(P, Q) = \mathrm{QF}^A(P, Q)$ and if I is the identity matrix, $\mathrm{QC}_{0.5}^I(P, Q) = \sqrt{2\chi^2(P, Q)}$.

3.2   Metric Properties

There are three conditions for a distance function, $\mathcal{D}$, to be a semi-metric. The first is *non-negativity* (*i.e.* $\mathcal{D}(P, Q) \geq 0$), the second is *symmetry* (*i.e.* $\mathcal{D}(P, Q) = \mathcal{D}(Q, P)$)

Algorithm 1 Quadratic-Chi Matlab Code for Full Histograms

```
function dist= QC(P,Q,A,m)

Z= (P+Q)*A;
% 1 can be any number as Z_i==0 iff D_i=0
Z(Z==0)= 1;
Z= Z.^m;
D= (P-Q)./Z;
% max is redundant if A is positive-semidefinite
dist= sqrt( max(D*A*D',0) );
```
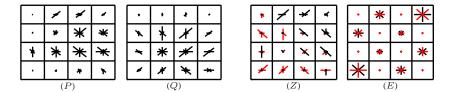

Fig. 2. This figure illustrates why it is not likely to get negative values in the square root argument of a QC distance for natural histograms and a typical similarity matrix. P and Q are two SIFT histograms. Z is the normalized difference vector. That is: $Z_i = \frac{P_i}{(\sum_c (P_c + Q_c) A_{ci})^m} - \frac{Q_i}{(\sum_c (P_c + Q_c) A_{ci})^m}$. Negative values are represented with red, positive values are represented with black. E is one of the eigenvectors of the similarity matrix that we used in the experiments which correspond to a negative eigenvalue. Z is very smooth while E is very non-smooth. This is typical of eigenvectors with negative values with typical parameters.

and the third is *subadditivity* (*i.e.* $\mathcal{D}(P, Q) \leq \mathcal{D}(P, K) + \mathcal{D}(K, Q)$). $\mathcal{D}$ is a metric if it is a semi-metric and it also has the property of *identity of indiscernibles* (*i.e.* $\mathcal{D}(P, Q) = 0$ if and only if $P = Q$).

A QC_m^A distance without the square root, is non-negative if the bin-similarity matrix, A, is positive-semidefinite. If A is positive-definitive, then it also has the property of *identity of indiscernibles*. This follows directly from the fact that the argument inside the square root in a QC histogram distance is a quadratic-form between two vectors. A QC histogram distance is symmetric if the bin-similarity matrix, A, is symmetric.

We now discuss subadditivity (*i.e.* $\mathcal{D}(P, Q) \leq \mathcal{D}(P, K) + \mathcal{D}(K, Q)$) for several distances. The χ^2 histogram distance is not subadditive. For example let $i = 0$, $k = 1$, $j = 2$ we get $\chi^2(i, j) = 1 > \chi^2(i, k) + \chi^2(k, j) = \frac{2}{3}$. However, $\sqrt{\chi^2}$ is subadditive for one and two dimensional non-negative histograms (verified by analysis). Experimentally it appears that $\sqrt{\chi^2}$ is subadditive for an N-dimensional non-negative histograms. Experimentally, QC members with the identity matrix seems to be subadditive for non-negative histograms. However, QC members with some positive-definitive bin-similarity matrices are not subadditive. The question when the QC histogram distances are subadditive is currently unresolved. An additional discussion about triangle inequality can be found in Jacobs et al. [21].

4   Cross-Bin Histogram Distance Properties

4.1   The Similarity-Matrix-Quantization-Invariance Property

The *Similarity-Matrix-Quantization-Invariance* property ensures that if two bins in the histograms have been erroneously quantized, this will not affect the distance. Mathematically we define this as:

Definition 1. *Let $\mathcal{D}$ be a cross-bin histogram distance between two histograms P and Q and let A be the bin-similarity/distance matrix. We assume P, Q and A are non-negative and that A is symmetric. Let $A_{k,:}$ be the kth row of A. Let $V = [V_1, \ldots, V_N]$*

be a non-negative vector and $0 \leq \alpha \leq 1$. *We define* $V^{\alpha,k,b} = [\ldots, \alpha V_k, \ldots, V_b + (1 - \alpha)V_k, \ldots]$. *That is,* $V^{\alpha,k,b}$ *is a transformation of V where* $(1 - \alpha)V_k$ *mass has moved from bin k to bin b. We define* $\mathcal{D}$ *to be* Similarity-Matrix-Quantization-Invariant *if:*

$$A_{k,:} = A_{b,:} \Rightarrow \forall \ 0 \leq \alpha \leq 1, \ 0 \leq \beta \leq 1 \ \mathcal{D}^A(P, Q) = \mathcal{D}^A(P^{\alpha,k,b}, Q^{\beta,k,b}) \tag{6}$$

We prove that EMD, $\widehat{\text{EMD}}$ and all the Quadratic-Chi histogram distances are *Similarity-Matrix-Quantization-Invariant* in the appendix [20].

4.2   The Sparseness-Invariance Property

The *Sparseness-Invariance* property ensures that distances between sparse histograms will be equal to distances between full histograms. Mathematically we define this as:

Definition 2. *Let* $\mathcal{D}$ *be a cross-bin histogram distance between two histograms* $P \in \mathcal{R}^N$ *and* $Q \in \mathcal{R}^N$ *and let A be the* $N \times N$ *bin similarity/distance matrix. Let* A' *be any* $(N + 1) \times (N + 1)$ *matrix whose upper-left sub-matrix equals A. We define* $\mathcal{D}$ *to be* Sparseness-Invariant *if:*

$$\mathcal{D}^A([P_1, \ldots, P_n], [Q_1, \ldots, Q_n]) = \mathcal{D}^{A'}([P_1, \ldots, P_n, \mathbf{0}], [Q_1, \ldots, Q_n, \mathbf{0}]) \tag{7}$$

QC members, EMD and the $\widehat{\text{EMD}}$ are *Sparseness-Invariant* directly from their definitions. A stronger property called *Extension-Invariance* was proposed by D'Agostino and Dardanoni for bin-to-bin distances [22]. This property requires that, if both histograms are extended by concatenating each of them with the same vector (not necessarily zeros), the distance is left unaltered. Cross-bin distances assumes dependence between histogram bins, thus this requirement is too strong for them.

4.3   Cross-Bin Histogram Distance Properties Discussion

A *Sparseness-Invariant* cross-bin histogram distance does not depend on the specific representation of the histograms (full or sparse). A *Similarity-Matrix-Quantization-Invariant* cross-bin histogram distance encompass its cross-bin relationships only in the bin-similarity matrix. Intuitively such properties are desirable. In the appendix [20], we compare experimentally distances which resembles QC distances, but are either not *Similarity-Matrix-Quantization-Invariant* or not *Sparseness-Invariant*. The comparison shows that these properties considerably boost performance (especially for sparse color histograms).

Rubner et al. [2,23] claim that one of the key advantages of the Earth Mover's Distance is that each compared object may be represented by an individual (possibly with a different number of bins) binning that is adapted to its specific distribution. The Quadratic-Form is regarded as not having this property (see for example, Table 1 in [23]). Since all the Quadratic-Chi histogram distances (including the Quadratic-Form) are both *Similarity-Matrix-Quantization-Invariant* and *Sparseness-Invariant* there is no obstacle to using them with individual binning; *i.e.* to use them to compare histograms that were adapted to each object individually.

Similarity-Matrix-Quantization-Invariant and *Sparness-Invariant* can contradict. For example, any distance applied to the transformed vectors $P'_i = \sum_c (P_c) A_{ci}$ and $Q'_i = \sum_c (Q_c) A_{ci}$ is *Similarity-Matrix-Quantization-Invariant*. However the χ^2 distance between P' and Q' is not *Sparseness-Invariant* (with respect to P and Q).

5 Implementation Notes

5.1 The Similarity Matrix and The Normalization Factor

It is desirable to have a transformation from a distance matrix into a similarity matrix, as many spaces are equipped with a useful distance (*e.g.* color space [24]). Hafner et al. [1] proposed this transformation:

$$A_{ij} = 1 - \frac{D_{ij}}{\max_{ij}(D_{ij})} \qquad (8)$$

Another possibility for choosing a similarity matrix is by using cross validation. However, we think that like for the Quadratic-Form, learning the similarity matrix (and for QC also the normalization factor) will be the best way to adjust them. This is left for future work. Currently we suggest to use thresholded ground distances as was used in [2,25,3,14] and choosing the normalization factor by cross validation.

5.2 Efficient Online Bin-Similarity Matrix Computation

For a fixed histogram configuration (*e.g.* SIFT, SC and IDSC) the bin-similarity matrix can be pre-computed once. Then, each distance computation is linear in the number of non-zero entries in the bin-similarity matrix.

There are cases where the bin-similarity matrix can not be pre-computed. For example, in our color experiments (Section 6.1), we used $N \times M$ color images as sparse histograms. That is, the query histogram was: $[1, \ldots, 1, 0, \ldots, 0]$ and each image being compared to the query was represented by the histogram: $[0, \ldots, 0, 1, \ldots, 1]$. Note that the full histogram dimension is $M \times N \times 256^3$, computing an $(M \times N \times 256^3)^2$ similarity matrix offline is not feasible. We can compute the similarity online for each pair of sparse histograms in $O((NM)^2)$ time. We now discuss how to do it more efficiently.

If we are comparing two images (as in Section 6.1) we can use a similarity matrix that gives far-away pixels zero similarity (see Eq. 10). Then, we can simply compare each pixel in one image to its corresponding $T \times T$ spatial neighbors in the second image. This reduces running time to $O(NMT^2)$. Using this technique, it is important to use a sparse representation for the bin-similarity matrix.

6 Results

We present results using the newly defined distances and state of the art distances, for image retrieval using SIFT-like descriptors and color image descriptors. In addition, we present results for shape classification using Inner Distance Shape Context (IDSC). More results for shape classification using SC, can be found in the appendix [20].

6.1 Image Retrieval Results

In this section we present results for image retrieval using the same benchmark as Pele and Werman [14]. We employed a database that contained 773 landscape images from

the COREL database that were also used in Wang et al. [26]. The dataset has 10 classes[1]:
People in Africa, Beaches, Outdoor Buildings, Buses, Dinosaurs, Elephants, Flowers,
Horses, Mountains and Food. The number of images in each class ranges from 50 to
100. From each class we selected 5 images as query images (images $1, 10, \ldots, 40$).
Then we searched for the 50 nearest neighbors for each query image. We computed the
distance of each image to the query image and its reflection and took the minimum. We
present results for two types of image representations: SIFT-like descriptors and small
L*a*b* images.

SIFT-like Descriptors The first representation - SIFT is a $6 \times 8 \times 8$ SIFT descriptor
[13] computed globally on the whole image. The second representation - CSIFT is a
SIFT-like descriptor on a color-edge image. See [14] for more details.

   We experimented with two new types of QC distances. The first is $QC_{0.5}^A$, which is
a cross-bin generalization of $\sqrt{2\chi^2}$, which we call Quadratic-Chi-Squared (QCS). The
second is $QC_{0.9}^A$, which has a larger normalization factor, which we call Quadratic-Chi-
Normalized (QCN). We do not use QC_m^A with $m \geq 1$ due to discontinuity problems,
see appendix [20] (practically, QC_1^A had slightly poorer results compared to $QC_{0.9}^A$).
We also experimented with the Quadratic-Form (QF) distance which is QC_0^A. For all
of these distances we used the bin-similarity matrix in Eq. 8. Let $M = 8$ be the num-
ber of orientation bins, as in Pele and Werman [14], the ground distance between bins
(x_i, y_i, o_i) and (x_j, y_j, o_j) is:

$$d_T(i,j) = \min\left(\left(||(x_i, y_i) - (x_j, y_j)||_2 + \min(|o_i - o_j|, M - |o_i - o_j|)\right) , T\right) \quad (9)$$

 We also used the identity matrix as a similarity matrix for all the above distances. We
also compared to L_2 and χ^2. $QF^I = L_2$, and nearest neighbors of χ^2 and QCS^I are the
same.

   We also compared to four EMD variants. The first was $\widehat{EMD}_1^D$ with $D = d_T$ (Eq.
9) as in Pele and Werman [3]. The second was the L_1 norm which is equal to $\widehat{EMD}_{0.5}^D$
with D equals to the Kronecker delta multiplied by two. The third is $SIFT_{DIST}$[3] which is
the sum of $\widehat{EMD}$ over all the spatial cells (each spatial cell contains one orientation his-
togram). The ground distance for the orientation histograms is: $\min(|o_i - o_j|, M - |o_i -
o_j|, 2)$ (M is the number of orientation bins). The fourth was the EMD-L_1[15] which is
EMD with L_1 as the ground distance. We also tried non-thresholded ground distances
(which produce non-sparse similarity matrices). However, the results were poor. This
is in line with Pele and Werman's findings that cross-bin distances should be used with
thresholded ground distances [14]. Finally, we compared to the Diffusion distance pro-
posed by Ling and Okada [16] and to three probabilistic based distances: Bhattacharyya
[17], Kullback-Leibler (KL) [18] and Jensen-Shannon (JS) [19] (we added Matlab's ep-
silon to all histogram bins when computing KL and JS throughout the paper, as they are
not well defined if there is a zero bin, without doing so accuracy was very low).

---

[1] The original database contains some visually ambiguous classes such as Africa that also con-
tains images of beaches in Africa. We used the filtered image dataset that was downloaded
from: http://www.cs.huji.ac.il/~ofirpele/FastEMD/

For each distance measure, we present the descriptor (SIFT/CSIFT) with which it performed best. The results for all the pairs of descriptors and distance measures can be found in the appendix[20]. The results are presented in Fig. 3(a) and show that $\text{QCN}^{1-\frac{d_{T=2}}{2}}$ (QCN with the similarity matrix: $A_{ij} = 1 - \frac{d_{T=2}(i,j)}{2}$) outperformed all other methods. $\widehat{\text{EMD}}_1^{d_{T=2}}$ ranked second. The computation of $\text{QCN}^{1-\frac{d_{T=2}}{2}}$ was 266 times faster than $\widehat{\text{EMD}}_1^{d_{T=2}}$, see Table 2 in page 12. QCN^I ranked third, which shows the importance of the normalization factor.

All cross-bin distances that use thresholded ground distances outperformed their bin-by-bin versions. The figure also shows that χ^2 and QF improve upon L_2. QCN and QCS which are mathematically sound combinations of χ^2 and QF outperformed both.

L*a*b* Images Our second type of image representation is a small L*a*b* image. We resized each image to 32×48 and converted them to L*a*b* space. The state of the art color distance is Δ_{00} - CIEDE2000 on L*a*b* color space[24,27]. As it is meaningful only for small distances we threshold it (as in [2,25,14]).

Again, we experimented with QCS, QCN and QF distances using the bin-similarity matrix in Eq. 8. The ground distance between two pixels $(x_i, y_i, L_i, a_i, b_i)$, $(x_j, y_j, L_j, a_j, b_j)$:

$$\text{s}(i,j) = ||(x_i, y_i) - (x_j, y_j)||_2$$

$$\text{dc}_{T_1,T_2}(i,j) = \begin{cases} \min\left((\text{s}(i,j) + \Delta_{00}((L_i, a_i, b_i), (L_j, a_j, b_j))), T_1\right) & \text{if s}(i,j) \leq T_2 \\ T_1 & \text{otherwise} \end{cases} \quad (10)$$

This distance is similar to the one used by [14], except that distances with spatial difference larger than the threshold T_2 are set to the maximum threshold T_1. This was done to accelerate the online computation of the bin-similarity matrix. The accuracy using this distance is the same as using the distance from Pele and Werman [14]. See appendix [20]. We also used $\widehat{\text{EMD}}$ with dc_{T_1,T_2} (Eq. 10) as a ground distance. Let I_1, I_2 be the two L*a*b* images. We also used the following distances:

$$L_1\Delta_{00} = \sum_{x,y} (\Delta_{00}(I_1(x,y), I_2(x,y))) \qquad L_1\Delta_{00}^T = \sum_{x,y} (\min(\Delta_{00}(I_1(x,y), I_2(x,y)), T))$$

$$L_2\Delta_{00} = \sum_{x,y} (\Delta_{00}(I_1(x,y), I_2(x,y)))^2 \qquad L_2\Delta_{00}^T = \sum_{x,y} (\min(\Delta_{00}(I_1(x,y), I_2(x,y)), T))^2$$

QCN^I, χ^2, L_2, L_1, $\text{SIFT}_{\text{DIST}}$[3], EMD-$L_1$[15], the Diffusion[16], Bhattacharyya [17], KL [18] and JS [19] distances cannot be applied to L*a*b* images as they are either bin-to-bin distances or applicable only to Manhattan networks.

We present results in Fig. 3. As shown, $\text{QCS}^{1-\frac{\text{dc}_{T_1=20,T_2=5}}{20}}$ and $\widehat{\text{EMD}}_1^{\text{dc}_{T_1=20,T_2=5}}$ [14] distances ranked first. $\text{QCS}^{1-\frac{\text{dc}_{T_1=20,T_2=5}}{20}}$ ran 300 times faster (see Table 2). However, since the computation of the bin-similarity matrix cannot be offline here, the real gain is a factor of 17. The $\text{QF}^{1-\frac{\text{dc}_{T_1=20,T_2=5}}{20}}$ distance ranked last, which shows the importance of the normalization factor of the QC histogram members.

Although a QC distance alleviates quantization problems, EMD does it better, instead of matching everything to everything it finds the optimal matching. EMD however,
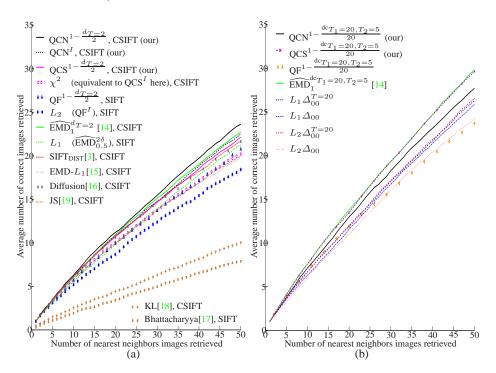
Fig. 3. Results for image retrieval.

(a) **SIFT-like descriptors.** For each distance measure, we present the descriptor (SIFT/CSIFT) with which it performed best. The results for all the pairs of descriptors and distance measures can be found in the appendix[20]. There are several key observations. First, the QC members performance is excellent. $QCN^{1-\frac{d_{T=2}}{2}}$ (QCN with the similarity matrix: $A_{ij} = 1 - \frac{d_{T=2}(i,j)}{2}$) outperformed all other distances. $\widehat{EMD}_1^{d_{T=2}}$ ranked second, but its computation was 266 times slower than $QCN^{1-\frac{d_{T=2}}{2}}$ computation (see Table 2). Second, all cross-bin versions of the distances (with d_T or a transformation of it) performed better than their bin-by-bin versions (with the identity matrix or the Kronecker delta function). Third, QCN^I ranked third, although its a bin-to-bin distance. This shows the importance of the normalization factor. Finally, χ^2 and QF improve upon L_2. However, χ^2 does not take cross-bin relationships into account and QF does not reduce the effect of large bins. QCS and QCN histogram distances, which are mathematically sound combinations of χ^2 and QF have the two properties and outperformed both.

(b) **L*a*b* images results.** QCN^I, χ^2, L_2, L_1, $SIFT_{DIST}$[3], EMD-L_1[15], Diffusion[16], Bhattacharyya[17], KL[18] and JS[19] distances are not applicable here. $QCS^{1-\frac{dc_{T_1=20,T_2=5}}{20}}$ and $\widehat{EMD}_1^{dc_{T_1=20,T_2=5}}$ [14] ranked first. $QCS^{1-\frac{dc_{T_1=20,T_2=5}}{20}}$ computation is 300 times faster than $\widehat{EMD}_1^{dc_{T_1=20,T_2=5}}$ without taking the bin-similarity matrix computation into account and 17 times faster when it is taken into account (see Table 2). $QF^{1-\frac{dc_{T_1=20,T_2=5}}{20}}$ ranked last, which shows the importance of the normalization factor in QC members.

does not reduce the effect of large bins. We conjecture that a variant of EMD which will reduce the effect of large bins will have an excellent performance.

6.2   Shape Classification Results

In this section we present results for shape classification using the same framework as Ling et al. [11,15,28]. We test for shape classification with the Inner Distance Shape Context (IDSC) [11]. The original Shape Context (SC) descriptor was proposed by Belongie et al. [10]. Belongie et al. [10] and Ling and Jacobs [11] used the χ^2 distance for comparing shape context histograms. Ling and Okada [15] showed that replacing χ^2 with EMD-L_1 improves results. We show that QC members yields the best results.

   We tested on the articulated shape data set [11,28], that contains 40 images from 8 different objects. Each object has 5 images articulated to different degrees. The dataset is very challenging because of the similarity between different objects. The original SC had a very poor performance on this dataset, see appendix [20].

   Again, we experimented with QCS, QCN and QF distances with the bin-similarity matrix in Eq. 8. The ground distance between two bins (r_i, o_i), (r_i, o_i) was (M is the number of orientation bins):

$$\mathrm{dsc}_T(i, j) = \min\left((|d_i - d_j| + \min(|o_i - o_j|, M - |o_i - o_j|), T\right) \tag{11}$$

We also used the identity matrix as a similarity matrix, and thus we also compare to L_2. χ^2 and QCS^I distances are not equivalent here as the distance is not used for nearest neighbors. We refer the reader to Belongie et al. paper to see its usage [10]. Practically, QCS^I slightly outperformed χ^2 in this task, see Table 1.

   We also compared to four EMD variants: $\widehat{\mathrm{EMD}}_1^D$ with $D = \mathrm{dsc}_T$ (Eq. 11), the L_1 norm, $\mathrm{SIFT}_{\mathrm{DIST}}$[3] and EMD-$L_1$[15]. Finally, we compared to the Diffusion distance proposed by Ling and Okada [16] and to three probabilistic based distances: Bhattacharyya [17], Kullback-Leibler (KL) [18] and Jensen-Shannon (JS) [19].

   To evaluate results, for each image, the four most similar matches are chosen from other images in the dataset. The retrieval result is summarized as the number of 1st, 2nd, 3rd and 4th most similar matches that come from the correct object. Table 1 shows the retrieval results. The $\mathrm{QCN}^{1-\frac{\mathrm{dsc}_{T=2}}{2}}$ outperformed all the other methods. QCN^I performance is again excellent, which shows the importance of the normalization factor.

	Top 1	Top 2	Top 3	Top 4	AUC%		Top 1	Top 2	Top 3	Top 4	AUC%
$\mathrm{QCN}^{1-\frac{\mathrm{dsc}_{T=2}}{2}}$	39	38	38	34	0.950	$\widehat{\mathrm{EMD}}_1^{\mathrm{dsc}_T=2}$	39	36	35	27	0.902
QCN^I	40	37	36	33	0.940	L_1	39	35	35	25	0.890
$\mathrm{QCS}^{1-\frac{\mathrm{dsc}_{T=2}}{2}}$	39	35	38	28	0.912	$\mathrm{SIFT}_{\mathrm{DIST}}$[3]	38	37	27	22	0.848
QCS^I	40	34	37	27	0.907	EMD-L_1[15]	39	35	38	30	0.917
χ^2	40	36	36	21	0.902	Diffusion[16]	39	35	34	23	0.880
$QF^{1-\frac{\mathrm{dsc}_{T=2}}{2}}$	40	34	39	19	0.897	Bhattacharyya[17]	40	37	32	23	0.895
L_2	39	35	35	18	0.873	KL[18]	40	38	36	29	0.938
						JS[19]	40	35	37	21	0.900

Table 1. Shape classification results. $\mathrm{QCN}^{1-\frac{\mathrm{dsc}_{T=2}}{2}}$ outperformed all other distances.

Again all cross-bin distances outperformed their bin-by-bin versions. Again, χ^2 and QF improved upon L_2. QCN and QCS which are mathematically sound combinations of χ^2 and QF outperformed both.

6.3 Running Time Results

All runs were conducted on a Pentium 2.8GHz. A comparison of the practical running time of all distances is given in Table 2. Clearly QCN and QCS distances are fast to compute. This is consistent with their linear time complexity. The only non-linear time distances are $\widehat{\text{EMD}}$ [14] and EMD-L_1[15] which are also practically much slower than the other methods. Our method can be easily parallelized, taking advantage of multi-core computers or the GPU.

Descriptor	QCN^{A_2}	QCN^I	QCS^{A_2}	QCS^I	χ^2	QF^{A_2}	L_2	$\widehat{\text{EMD}}^{D_2}$[14]	L_1	$SIFT_{DIST}$[3]
(SIFT)	**0.15**	**0.1**	**0.07**	**0.014**	**0.013**	**0.05**	**0.011**	**40**	**0.011**	**0.07**
(IDSC)	**6.41**	**2.99**	**2.32**	**0.35**	**0.34**	**1.25**	**0.14**	**133.75**	**0.32**	**0.31**

Descriptor	EMD-L_1[15]	Diffusion[16]	JS[19]	KL[18]	Bhattacharyya[17]
(SIFT)	**40**	**0.27**	**0.088**	**0.048**	**0.015**
(IDSC)	**20.57**	**3.15**	**1.40**	**8.53**	**17.17**

Descriptor	$QCN^{A_{20}}$	$QCS^{A_{20}}$	$QF^{A_{20}}$	$\widehat{\text{EMD}}^{D_{20}}$[14]	$L_1\Delta_{00}^{T=20}$	$L_1\Delta_{00}$	$L_2\Delta_{00}^{T=20}$	$L_2\Delta_{00}$
(L*a*b*)	**20** (370)	**19** (369)	**11** (361)	**6000** (6350)	**3.2**	**3.2**	**3.2**	**3.2**

Table 2. (SIFT) 384-dimensional SIFT-like descriptors matching time (in *milliseconds*). The distances from left to right are the same as the distances in Fig. 3 (a) from up to down.
(IDSC) 60-dimensional IDSC histograms matching time (in *microseconds*). The distances from left to right are the same as the distances in Table 1 from up to down.
(L*a*b*) 32×48 L*a*b* images matching time (in *milliseconds*). The distances from left to right are the same as the distances in Fig. 3 (b) from up to down. In parentheses is the time it takes to compute the distance and the bin-similarity matrix as it cannot be computed offline.

7   Conclusions

We presented a new cross-bin distance family - the Quadratic-Chi (QC). QC distances have many desirable properties. Like the Quadratic-Form histogram distance they take into account cross-bin relationships. Like χ^2 they reduce the effect of large bins. We formalized two new cross-bin properties, *Similarity-Matrix-Quantization-Invariance* and *Sparseness -Invariance*. QC members were shown to have both. Finally, QC distance

computation time is linear in the number of non-zero entries in the bin-similarity matrix. Experimentally, QC outperformed state of the art distances, while having a very short run-time.

There are several open questions that we still need to explore. The first is for which QC distances does the the triangle inequality holds for. The second is whether we can change the Earth Mover's Distance so that it will also reduce the effect of large bins. Concave-cost network flow [29] seems to be the right direction for future work although it presents two major obstacles. First, the concave-cost network flow optimization is NP-hard [29]. However, there are available approximations [29,30]. Second, simply using concave-cost flow networks will result in a distance which is not *Similarity-Matrix-Quantization-Invariant*. We would also like to explore whether metric learning methods such as [31,32,33,34,35,36,37,38] can be generalized for the Quadratic-Chi histogram distance. Assent et al. [39] have suggested methods that accelerate database retrieval that uses Quadratic-Form distances. Generalizing these methods for the Quadratic-Chi distances is of interest. Finally, other computer vision applications such as tracking can use the QC distances. The project homepage, including code (C++ and Matlab wrappers) is at: http://www.cs.huji.ac.il/~ofirpele/QC/.

References

1. Hafner, J., Sawhney, H., Equitz, W., Flickner, M., Niblack, W.: Efficient color histogram indexing for quadratic form distance functions. PAMI (1995) 1, 7
2. Rubner, Y., Tomasi, C., Guibas, L.J.: The earth mover's distance as a metric for image retrieval. IJCV (2000) 2, 6, 7, 9
3. Pele, O., Werman, M.: A linear time histogram metric for improved sift matching. In: ECCV. (2008) 2, 3, 7, 8, 9, 10, 11, 12
4. Snedecor, G., Cochran, W.: Statistical Methods, ed 6. Ames, Iowa (1967) 3
5. Cula, O., Dana, K.: 3D texture recognition using bidirectional feature histograms. IJCV (2004) 3
6. Zhang, J., Marszalek, M., Lazebnik, S., Schmid, C.: Local features and kernels for classification of texture and object categories: A comprehensive study. IJCV (2007) 3
7. Varma, M., Zisserman, A.: A statistical approach to material classification using image patch exemplars. PAMI (2009) 3
8. Xu, D., Cham, T., Yan, S., Duan, L., Chang, S.: Near Duplicate Identification with Spatially Aligned Pyramid Matching. CSVT (accepted) 3
9. Forssén, P., Lowe, D.: Shape Descriptors for Maximally Stable Extremal Regions. In: ICCV. (2007) 3
10. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. PAMI (2002) 3, 11
11. Ling, H., Jacobs, D.: Shape classification using the inner-distance. PAMI (2007) 3, 11
12. Martin, D., Fowlkes, C., Malik, J.: Learning to detect natural image boundaries using local brightness, color, and texture cues. PAMI (2004) 3
13. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. IJCV (2004) 3, 8
14. Pele, O., Werman, M.: Fast and robust earth mover's distances. In: ICCV. (2009) 3, 7, 8, 9, 10, 12
15. Ling, H., Okada, K.: An Efficient Earth Mover's Distance Algorithm for Robust Histogram Comparison. PAMI (2007) 3, 8, 9, 10, 11, 12

16. Ling, H., Okada, K.: Diffusion distance for histogram comparison. In: CVPR. (2006) 3, 8, 9, 10, 11, 12
17. Bhattacharyya, A.: On a measure of divergence between two statistical populations defined by their probability distributions. BCMS (1943) 3, 8, 9, 10, 11, 12
18. Kullback, S., Leibler, R.: On information and sufficiency. AMS (1951) 3, 8, 9, 10, 11, 12
19. Lin, J.: Divergence measures based on the Shannon entropy. IT (1991) 3, 8, 9, 10, 11, 12
20. Pele, O., Werman, M.: The quadratic-chi histogram distance family - appendices. http://www.cs.huji.ac.il/~ofirpele/publications/ECCV2010app.pdf (2010) 4, 6, 7, 8, 9, 10, 11
21. Jacobs, D., Weinshall, D., Gdalyahu, Y.: Classification with nonmetric distances: Image retrieval and class representation. PAMI (2000) 5
22. D'Agostino, M., Dardanoni, V.: Whats so special about Euclidean distance? SCW (2009) 6
23. Rubner, Y., Puzicha, J., Tomasi, C., Buhmann, J.: Empirical evaluation of dissimilarity measures for color and texture. CVIU (2001) 6
24. Luo, M., Cui, G., Rigg, B.: The Development of the CIE 2000 Colour-Difference Formula: CIEDE2000. CRA (2001) 7, 9
25. Ruzon, M., Tomasi, C.: Edge, Junction, and Corner Detection Using Color Distributions. PAMI (2001) 7, 9
26. Wang, J., Li, J., Wiederhold, G.: SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture LIbraries. PAMI (2001) 8
27. Sharma, G., Wu, W., Dalal, E.: The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations. CRA (2005) 9
28. Ling, H.: Articulated shape benchmark and idsc code. http://www.ist.temple.edu/~hbling/code/inner-dist-articu-distrbution.zip (2010) 11
29. Guisewite, G., Pardalos, P.: Minimum concave-cost network flow problems: Applications, complexity, and algorithms. AOR (1990) 13
30. Amiri, A., Pirkul, H.: New formulation and relaxation to solve a concave-cost network flow problem. JORS (1997) 13
31. Xing, E., Ng, A., Jordan, M., Russell, S.: Distance metric learning with application to clustering with side-information. NIPS (2003) 13
32. Bar-Hillel, A., Hertz, T., Shental, N., Weinshall, D.: Learning distance functions using equivalence relations. In: ICML. (2003) 13
33. Goldberger, J., Roweis, S., Hinton, G., Salakhutdinov, R.: Neighbourhood components analysis. NIPS (2005) 13
34. Globerson, A., Roweis, S.: Metric learning by collapsing classes. NIPS (2006) 13
35. Yang, L., Jin, R.: Distance metric learning: A comprehensive survey. MSU (2006) 13
36. Davis, J., Kulis, B., Jain, P., Sra, S., Dhillon, I.: Information-theoretic metric learning. In: ICML. (2007) 13
37. Yu, J., Amores, J., Sebe, N., Radeva, P., Tian, Q.: Distance learning for similarity estimation. PAMI (2008) 13
38. Weinberger, K., Saul, L.: Distance metric learning for large margin nearest neighbor classification. JMLR (2009) 13
39. Assent, I., Wichterich, M., Seidl, T.: Adaptable Distance Functions for Similarity-based Multimedia Retrieval. DSN (2006) 13