TOWARD AN IMPROVED ERROR METRIC

Qi Tian, Qing Xue, Jie Yu, Nicu Sebe[1], *Thomas S. Huang*[2]

Department of Computer Science, University of Texas at San Antonio, TX 78249
[1]Faculty of Science, University of Amsterdam, The Netherlands
[2]Beckman Institute, University of Illinois, Urbana, IL 61801

ABSTRACT

In many computer vision algorithms, the well known Euclidean or SSD (sum of the squared differences) metric is prevalent and justified from a maximum likelihood perspective when the additive noise is Gaussian. However, Gaussian noise distribution assumption is often invalid. Previous research has found that other metrics such as double exponential metric or Cauchy metric provide better results, in accordance with the maximum likelihood approach. In this paper, we examine different error metrics and provide a theoretical approach to derive a rich set of nonlinear estimations. Our results on image databases show more robust results are obtained for noise estimation based on the proposed error metric analysis.

1. INTRODUCTION

In computer vision community, a metric or similarity measure is used to determine the distance between two features. The SSD (L_2 -sum of the squared differences) is the most commonly used metric. But it has been suggested that this metric is not appropriate for many problems [1]. The SAD (L_1 -sum of the absolute differences) is another commonly used similarity metric. From a maximum likelihood perspective, it is well known that the SSD is justified when the additive noise distribution is Gaussian [2]. The SAD is justified when the additive noise distribution is Exponential (double or two-sided exponential) [2]. The common assumption is that the real noise distribution should fit either the Gaussian or the Exponential. But in practice, Gaussian or Exponential assumption is often invalid.

Content-based image retrieval is a fast-growing research area in the recent years. In general, image retrieval by content requires algorithms for extracting and computing features. Extracted features from the imagery may be associated with entire digital images, or perhaps with specific region of interest. A method for calculating the similarity between two digital images is to compare its Euclidean distance between their feature vectors. In previous work, most of the attention has been focused on the low-level feature model such as color, texture, and shape with little or no consideration of the noise models.

Sebe *et al* [2, 3] addressed the noise model problem from the maximum likelihood point of view. In three different applications of content-based image retrieval, stereo matching, and motion tracking, they experimented with different modeling functions for the noise distribution. They found the Cauchy metric fits the real noise distribution better than the Gaussian or the Exponential models.

Although this work is one of the first (if not first) efforts to address the noise model in similarity image retrieval, the connection between noise model and the error metric has not been fully studied yet. In [2], the noise distributions such as Gaussian, Exponential, and Cauchy result in L_2 , L_1 or L_c metrics. This arises the question: what about other possible connections between the noise model and error metric? whether we could find a more accurate noise model than the noise models mentioned above based on the error metric analysis. It is this problem that we examined in this paper. We proposed a rich set of error metrics besides the L_2 , L_1 , and L_c , and derived the corresponding nonlinear estimations such as harmonic mean, geometric mean, as well as their generalized nonlinear operations. It not only achieves a better noise estimation than the conventional L_2 and L_1 metrics, but also provides a possible way to disclose the connection between the noise model and the corresponding error metric.

2. ERROR METRIC ANALYSIS

2.1 Maximum likelihood approach

The additive noise model is the dominant model used in computer vision regarding maximum likelihood estimation. Haralick and Shapiro [4] consider this model in defining the M-estimate: "Any estimate μ defined by a

0-7803-8554-3/04/$20.00 ©2004 IEEE.

minimization problem of the form $\min \sum_i f(x_i - \mu)$ is called an M-estimate." Note that the operation "–" between the estimate and the real data implies an additive model.

Maximum likelihood theory [2, 3, 7] allows us to relate a noise distribution to an error metric. Specifically, if we are given the noise distribution, then the error metric which maximizes the similarity probability [2, 3] is

$$\sum_{i=1}^{M} \rho(n_i) \qquad (1)$$

where n_i represents the i^{th} bin of the discrete noise and ρ is the maximum likelihood estimate of the negative logarithm of the probability density function of the noise. In practice, the noise is typically represented by the difference between the corresponding elements of feature vectors.

In [2], *maximum likelihood* gives a direct connection between the noise distribution and the comparison metrics. One can note that the Gaussian model is related to L_2 metric, while the Exponential model is related to L_1 metric, as well as Cauchy model metric, respectively.

2.2 Error metric analysis

As we mentioned above, the mathematical connection was shown in [2, 3] between the closed-form noise distribution such as Gaussian, Exponential, or Cauchy and the corresponding error metrics. What are the other possible error metrics, and whether they could provide a better estimation than the ones above? We consider this problem as *error metric analysis*. We give the mathematical formulation in the following.

Suppose a set of observations

$$x_i = \mu + n_i \qquad (2)$$

are given, where μ is the original signal (to be estimated) and n_i , $i = 1, \cdots, N$ are zero mean noise components. Then for some function

$$f(x, \mu) \geq 0 \qquad (3)$$

which satisfies the condition $f(\mu, \mu) = 0$, an estimator $\hat{\mu}$ of μ can be defined which minimizes

$$\varepsilon = \sum_{i=1}^{N} f(x, \hat{\mu}) \qquad (4)$$

or which equivalently satisfies

$$\sum_{i=1}^{N} \frac{d}{d\hat{\mu}} f(x_i, \hat{\mu}) = 0 \qquad (5)$$

In some special cases, the closed-form solution $\hat{\mu} = g(x_1, x_2, \cdots, x_N)$ can be obtained. The arithmetic mean, median, harmonic mean, and geometric mean are within this category and shown in Table 1. Notice that the error metric for the L_2 metric (SSD) results in the arithmetic mean, and the L_1 metric (SAD) results in the median. However, no literature has discussed the harmonic mean and the geometric mean from the error metric point of view.

Table 1. Estimation based on different error metrics

	Estimation	Error Metric
Arithmetic mean	$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i$	$\varepsilon = \sum_{i=1}^{N} (x_i - \hat{\mu})^2$
Median	$\hat{\mu} = med(x_1, \cdots, x_N)$	$\varepsilon = \sum_{i=1}^{N} \lvert x_i - \hat{\mu} \rvert$
Harmonic mean	$\hat{\mu} = \frac{N}{\sum_{i=1}^{N} \frac{1}{x_i}}$	$\varepsilon = \sum_{i=1}^{N} x_i (\frac{\hat{\mu}}{x_i} - 1)^2$
Geometric mean	$\hat{\mu} = (\prod_{i=1}^{N} x_i)^{\frac{1}{N}}$	$\varepsilon = \sum_{i=1}^{N} [\log(\frac{x_i}{\hat{\mu}})]^2$

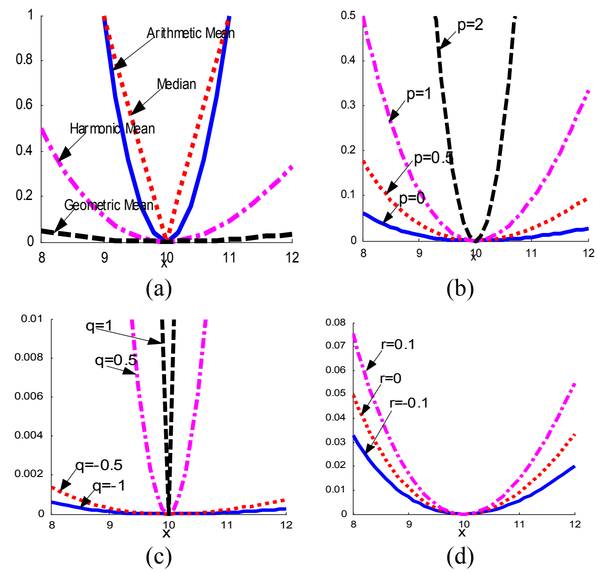


Figure 1. The error metric function $f(x, \mu)$ of (a) the arithmetic mean, median, harmonic mean and geometric mean and generalized harmonic mean (b) 1st-type, (c) 2nd-type, and (d) the generalized geometric mean (μ is fixed and set to 10)

If one infers the right forms of the error metrics, the proof of Table 1 is quite straightforward. Figure 1(a) shows the error metric function $f(x, \hat{\mu})$ for the arithmetic mean, median, harmonic mean and geometric mean, respectively, where the estimate (μ) is assumed to have the same value, e.g., 10 as an example in Fig. 1(a), for the purpose of comparison. Note that, for those nonlinear estimations concerned, the observations which are far away from the correct estimate (μ) will make less contribution to producing μ, as distinct from the arithmetic mean. In these cases the estimated values will be less sensitive to the bad observations (i.e., observation with large variance), and they are therefore more robust.

2.3 A generalization of error metric

In order to study further the robust property of different error metrics, the generalization of the harmonic mean and the geometric mean have been made as shown in Table 2. The proof of Table 2 is also straightforward using Equation (5).

Table 2. A generalization of error metric

	Estimation	Error Criteria
Generalized harmonic mean (1st type)	$\hat{\mu}=\frac{\sum_{i=1}^{N}(x_i)^{p-1}}{\sum_{i=1}^{N}(x_i)^{p-2}}$	$\varepsilon=\sum_{i=1}^{N}(x_i)^p(\frac{\hat{\mu}}{x_i}-1)^2$
Generalized harmonic mean (2nd type)	$\hat{\mu}=[\frac{N}{\sum_{i=1}^{N}(x_i)^q}]^{-\frac{1}{q}}$	$\varepsilon=\sum_{i=1}^{N}[(x_i)^q-(\hat{\mu})^q]^2$
Generalized geometric mean	$\hat{\mu}=[\prod_{i=1}^{N}(x_i)^{(x_i)^{2r}}]^{\frac{1}{\sum_{i=1}^{N}(x_i)^{2r}}}$	$\varepsilon=\sum_{i=1}^{N}[(x_i)^r\log(\frac{x_i}{\hat{\mu}})]^2$

Note that the parameters p, q, r in Table 2 can take any real numbers. In terms of the generalized harmonic mean estimation, it is interesting to point out that the 1st type is generalized based on the error metric representation, while the 2nd type is generalized based on the estimation representation. However, if $p=1$, $q=-1$, both types will become ordinary harmonic mean and if $p=2$ and $q=1$, both types will become arithmetic mean. As regards the generalized geometric mean estimation, if $r=0$, it will become ordinary geometric mean. Therefore, the generalization made here actually covers a wide range of operations. Figure 2 (b)-(d) show the error metrics for the 1st type and 2nd type generalized harmonic mean, and generalized geometric mean operations, respectively.

It should also be noted that Tables 1 and 2 show the closed-form estimation for some special cases. In some other cases there may not be closed-form solution of $\hat{\mu}$ in terms of x_i, $i=1,\cdots,N$. In this case, $\hat{\mu}$ can still be estimated by numerical analysis, i.e., greedy search of $\hat{\mu}$ to minimize ε.

3. NOISE ESTIMATION IN IMAGE DATABASE

We applied the theoretical results described in Section 2 to determine the influence of similarity error metric on the noise distribution estimation in the Corel database.

The Corel database is currently composed of 12000 images. The problem is to model the noise between two images, which are different due to different handling and storage conditions of the original photographs, varying orientation, motion, or printer noise. For our experiments, we randomly used 300 images from the Corel database as the ground truth dataset. For each selected image, a copy is printed, digitized on a scanner, and resized. The whole process introduces noise due to the dithering patterns of the printer and scanner. We repeated the process 10 times for each selected image. Therefore, we have total 3300 images including 300 original images and 10 copies for each original image. They represent the near-copies.

According to this ground truth, we determined the real distribution of the similarity noise considering two different spaces: image space (or intensity space) and feature space. In image space, the intensity of pixel is used. We formed a long column-wise vector by concatenating all the pixels in the image column by column. Since the image space is usually too large to handle (e.g., the average number of pixels per image is $500\times700=350{,}000$), we randomly choose a subset of pixels (e.g., 1% and 0.1% of all pixels) but we keep the pixel correspondence among all the testing images. In feature space, we have used two visual features: wavelet-based texture [6], and edge-based structure feature [7]. Color features cannot be used since after printing, the images became black and white.

As a comparison, we compared with the conventional L_2 and L_1 metrics, and also the Cauchy metric. Specifically, we want to point out that an estimate is regarded as *optimal* with respected to the specified error metric, e.g., L_2 in the least mean square sense. Unless the correct connection is found between the noise model and the error metric, the estimation of noise will not be optimal. For instance, L_2 metric is optimal for Gaussian noise estimation from the maximum likelihood point view, but L_2 metric is not optimal for Exponential, or Cauchy distributed noises. Therefore, which error metric is more appropriate depends on the real noise distribution. Unless we have the prior knowledge about the real noise distribution, we have to examine all the possible error metrics to find the best fit.

Our algorithm can be described as follows:

Step 1: Compute the feature vectors in the image space and visual feature space.

Step 2: Compute the real noise distribution from the difference between the corresponding elements of feature vectors of the original images and their copies.

Step 3: Estimate the noise under different error metrics as shown in Tables 1 and 2. The probability density function of the estimated noise can be obtained by normalizing the corresponding histogram.

Step 4: Compare the probability density function M of the estimated noise to the real noise distribution R using the Chi-square test [7].

$$\chi^2=\sum_i\frac{(R_i-M_i)^2}{M_i} \tag{5}$$

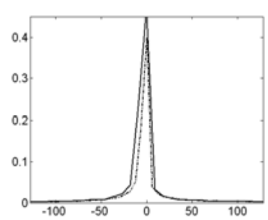
where the sum is over all bins.

Step 5: Select the corresponding error metric of the best fit to the real distribution.

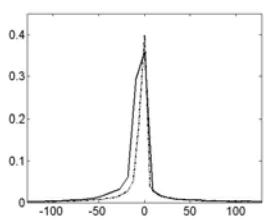
Table 3 shows the results of Chi-square test for various error metrics. The parameters of *p*, *q*, and *r* are tested in the range of –5 to 5 with step size 0.1. For Cauchy metric, $\varepsilon = \sum_{i=1}^{N} \log(1+\frac{(x_i-\hat{\mu})^2}{a^2})$, there is no closed-form solution for $\hat{\mu}$ as those shown in Tables 1 and 2. Instead, $\hat{\mu}$ can be estimated numerically by a greedy search to minimize ε in a wide range for both $\hat{\mu}$ and parameter *a*. The range of $\hat{\mu}$ can be centered at the arithmetic mean with a large window size. The range of parameter *a* can be centered at the value determined by best match between the real noise distribution and the one modeled by a Cauchy distribution similar as in [3].

Table 3. The Chi-square test values for error metrics

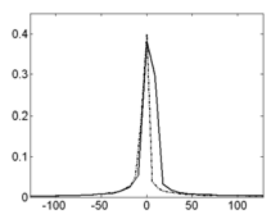
Error Metric	Structure	Texture	Intensity
L_1	0.2022	0.5269	2.6562
L_2	0.2425	0.5017	4.1066
Cauchy	0.3365 (*a*=5)	0.4953 (*a*=15.1)	3.0607 (*a*=24.8)
Harmonic	0.3087	0.2701	3.0970
Geometric	0.3138	0.4834	3.2940
1st type generalized harmonic (gh)	**0.1648** (*p*=1.8)	0.1679 (*p*=-4.9)	3.0970 (*p*=1.0)
2nd type generalized harmonic (gh)	0.2082 (*q*=-0.1)	**0.1331** (*q*=-0.7)	2.2317 (*q*=-3.0)
Generalized geometric (gg)	0.3086 (*r*=4.9)	0.2959 (*r*=4.7)	**2.0374** (*r*=-4.2)
best metric	1st gh (*p*=1.8)	2nd gh (*q*=-0.7)	gg (*r*=-4.2)



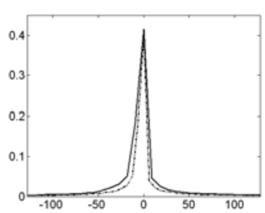
(a) L_1 (0.2022)



(b) L_2 (0.2425)



(c) Cauchy (0.3365)



(d) Generalized harmonic mean with *p*=1.8 (0.1648)

Figure 2. The real noise distribution (dashed line) vs. the estimated noise distribution (solid line) for structure feature

In Table 3, the smaller the Chi-square test value, the closer of the estimation is to the real distribution. There are several conclusions: (1) the L_1 metric and Cauchy metric are more suitable than L_2 metric, e.g., in the intensity (image) space. This observation agrees with the results in [2, 3]. (2) More accurate estimations can be obtained by a large set of other error metrics than by L_1, L_2, and Cauchy metrics, e.g., in the structure feature space. This shows the effectiveness of the proposed error metric analysis. Especially those nonlinear estimations based on the generalized harmonic mean and generalized geometric mean are more robust than those based on the L_2 metric and L_1 metric.

Figure 2 shows the probability density function of the real noise and the estimated noise by different error metrics. The Chi-square test value is shown for each metric. Clearly, the 1st type generalized harmonic mean with $p = 1.8$ gives the best fit among all error metrics.

4. SUMMARY

In this paper, the estimation of noise distribution in image databases can be significantly improved based on some new error metrics instead of the well known SSD and SAD metrics. Therefore, the error metric analysis can be used as a general guide for designing the most appropriate operations for the robust fusion problems. Our future work includes applying error metric analysis on stereo matching and motion tracking problems.

5. REFERENCES

[1] M. Zakai, "General Error Criteria," *IEEE Trans. on Information Theory*, pp. 94-95, January 1964.

[2] N. Sebe, M. S. Lew, and D. P. Huijsmans, "Toward Improved Ranking Metrics," *IEEE Trans. Pattern Analysis and Machine Intelligence*, pp. 1132-1143, Oct. 2000.

[3] N. Sebe, M. S. Lew, and D. P. Huijsman, "Which Ranking Metric is Optimal? with Applications in Image Retrieval and Stereo Matching," in *Proc. IEEE Int'l Conf. Pattern Recognition*, pp. 265-271, 1998.

[4] R. Haralick and L. Shapiro, *Computer and Robot Vision II*, Addison-Wesley, 1993.

[5] J. R. Smith, and S. F. Chang, "Transform Features for Texture Classification and Discrimination in Large Image Database," in *Proc. IEEE Intl. Conf. on Image Proc.*, 1994.

[6] S. X. Zhou, Y. Rui and T. S. Huang, "Water-filling algorithm: A novel way for image feature extraction based on edge maps," in *Proc. IEEE Intl. Conf. On Image Proc.*, Japan, 1999.

[7] P. J. Huber, *Robust Statistics*, New York: John Wiley & Sons, 1981.