Image fusion performance metric based on mutual information and entropy driven quadtree decomposition

M. Hossny, S. Nahavandi, D. Creighton and A. Bhatti

The mutual information (MI) measure has become a popular metric to assess image fusion performance. However, despite its publicity, it provides a questionable result that consistently favours additive fusion (averaging) over multi-scale decomposition (MSD) fusion algorithms. Presented is a localised variation of MI to assess image fusion performance while preserving the importance of local structural similarity. The presented metric has been validated with extensive tests on popular image fusion test cases.

Introduction: Image fusion performance metrics depend mainly on estimating the number of informative features transferred from both source images into the resulting fused image. The mutual information (MI) metric values the transfer of statistical properties from source images to the fused image [1, 2]. However, unlike gradient-based metrics [3, 4], it lacks in estimating the transfer of local structures from source images into the fused image. This limitation has led to a questionable result of MI consistently favouring simple averaging over multi-scale-decomposition (MSD) fusion algorithms [2]. This Letter introduces a localised mutual information (LMI) measure as a metric for image fusion performance assessment. The proposed metric values the structural similarity and gradient information as well as statistical properties that compensate for the luminance shift and contrast degradation problems.

Mutual information (MI): In information theory, mutual information measures the distance between a joint statistical distribution of two random variables X and Y from the case of being totally independent. In terms of entropies, and using Kvalseth normalisation [5], MI can be formulated as:

$$I(X,Y) = 2\frac{H(X)+H(Y)-H(X,Y)}{H(X)+H(Y)} \tag{1}$$

where $H(X)$, $H(Y)$, $H(X,Y)$ are the entropies of X, Y, and their joint probability, respectively. According to (1), $I(X,Y)$ measures the closeness between two normalised histograms X and Y. The result ranges between zero for completely independent image histograms and 1 for identical normalised histograms. Qu *et al.* [1] proposed using the mutual information measure to estimate the joint information between source images x, y and the fused image f as follows:

$$M_F^{XY} = I(F,X) + I(F,Y) \tag{2}$$

where X, Y and F are the normalised histograms of source images x, y and the fused image f, respectively. The main advantage of using MI as an image fusion metric is its immunity to the luminance change of the fused image. However, its main drawback is its lack of measuring structural information.

MI's disputable result: The problem with MI, as a fusion metric, was first highlighted by Qu *et al.* when they introduced MI as a performance metric for image fusion [1]. The results presented in their paper suggested that simple averaging is better than discrete wavelet transform fusion. They did not declare it as a problem and explained the odd result to be due to performing fusion at an unmatched number of levels of detail. Afterwards, Cvejic *et al.* [2] generalised MI favouring averaging results to be a common disadvantage of MI as a fusion performance metric. They highlighted the lack of tuning parameters in MI compared to gradient-based metrics [3] and [4] and suggested a new mutual information measure based on a tunable entropy such as Tsallis's [6].

The explanation of MI's questionable results is fairly simple. Gradient-based metrics value edges and high frequencies. Additionally, gradient-based metrics weigh gradient information in source images differently using Piella and Hiejman's saliency functions in [4]. On the other hand, MI measures the closeness between two image histograms regardless of how the information is spatially distributed across the images. This allows MI to be less sensitive to luminance shifting and contrast loss fusion artefacts of additive fusion as illustrated in Fig. 1. These artefacts were the main motivation behind applying fusion at different levels of detail.

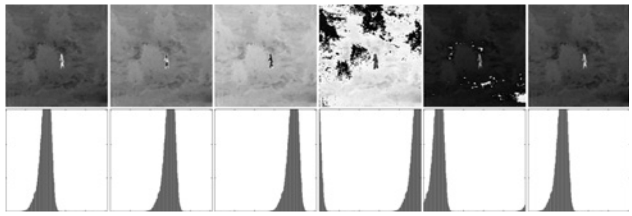

Fig. 1 *Increasing luminance of source image (far left) has no effect on mutual information. Left to right: Luminance increased by 48 intensity levels modulo number of grey levels (e.g. 256 for grey-scaled images). All these images are MI-similar*

Localised mutual information (LMI): The proposed metric applies the classic mutual information metric on local blocks in source and fused images. Localisation is achieved via entropy driven quadtree decomposition. Quadtree decomposition $Q(x)$ subdivides x into four equally sized quarters $\{x_1, x_2, x_3, x_4\}$ if the entropy $H(X) \geq \epsilon_x$. The algorithm then recursively applies the entropy examination and subdivision on each $x_{i|i=1\ldots4}$ as follows:

$$Q(x) = \begin{cases} \cup_{i=1}^{4} Q(x_i) & \text{if } H(X) \geq \epsilon_x \\ \{x\} & \text{otherwise} \end{cases} \tag{3}$$

where $H(X)$ is the entropy of an image (or possibly sub-image) x, capitalised X is the probability distribution of colours in x, and ϵ_x is the cutoff entropy. To control localisation of the MI the cutoff entropy was estimated as the average local entropies of blocks with a kernel-window of size w applied to the image being decomposed x as follows:

$$\epsilon_x^W = \frac{1}{|W|}\sum_{w\in W} H(X|w) \tag{4}$$

where w is the localisation kernel window, and $|W|$ is the number of all local blocks. The calculated cutoff entropy guarantees the decomposition of both images to a minimum block size d equal to the size of the kernel window w at which local entropies were calculated. The additional complexity this step imposes is upper-bounded with $O(N)$, where N is the number of pixels per kernel window. The overall quadtree decomposition with minimum block size parameter determining the cutoff entropy in (4) can then be described as follows:

$$Q^W(x) = \begin{cases} \cup_{i=1}^{4} Q^W(x_i) & \text{if } H(X) \geq \epsilon_x^W \\ \{x\} & \text{otherwise} \end{cases} \tag{5}$$

As there is no guarantee that entropy-similar images have the same information distribution across the spatial dimensions, the proposed algorithm decomposes an image and imposes the extracted topology on the other image to maintain similar block sizes and vice versa. The proposed metric can then be defined as follows:

$$I_{W,\alpha}^{Q}(x,y) = \ldots \tag{6}$$

$$\sum_{x_i\in Q^W(x)} \frac{d(x_i)^\alpha I(X_i,Y_i)}{2D_\alpha(x)} + \sum_{y_i\in Q^W(y)} \frac{d(y_i)^\alpha I(X_i,Y_i)}{2D_\alpha(y)} \tag{7}$$

where $\alpha \in [-1, 1]$ weights each examined block in both images, $d(\bullet)$ returns the size of the block being examined in images x and y, and $D_\alpha(\bullet)$ is the sum of all weighted block sizes across the decomposed image:

$$D_\alpha(x) = \sum_{x_i\in Q(x)} d(x_i)^\alpha \tag{8}$$

Experiments and results: The experiments presented in this Section ran extensive duality tests on normalised MI [7], gradient-based fusion metrics (QE) [3], and the proposed metric (LMI) described in (6) against averaging (AVG), wavelet (DWT), Laplacian (LAP), morphological (MOD), and contrast (CON) MSD-fusion algorithms. The experiments were conducted on the three widely popular TNO's image fusion test cases, Dune (46 images), Trees (38 images), and UN Camp (64 images) available in [8]. Fused images were obtained using source images from the data set and a single coloured image ($\mathbf{0}$–image) [9]. The $\mathbf{0}$–images were approximated using the average colour intensity of all images in each data set [9]. MSD-fusion algorithms performed fusion at 2, 3, 4, and 5 levels of detail, selected high frequencies using a 5×5 consistency verification kernel [10], and averaged the intensities at the lowest level of detail. LMI was tested using five different minimum block sizes $w = \{2, 4, 8, 16, 32\}$. MI, QE, and LMI were tested without weighting functions. Then LMI was tested using the three variations $\alpha = \{-1, 0, 1\}$. The whole experiments generated 50320 fusion test cases.

The first experiment aimed to validate the advantage of LMI over classic MI and QE. Fig. 2 illustrates the behaviour of LMI as the number of levels of details l at which fusion takes place increases. The results in Fig. 2 also show LMI preserves the MI's luminance invariance advantage by recording the lowest duality error for AVG, while following QE's decaying duality index as minimum block size d decreases.

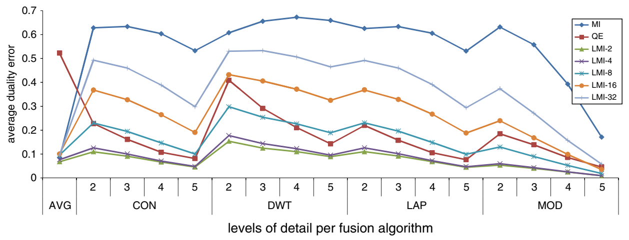

Fig. 2 *Average duality errors of MI, QE, and LMI fusion performance metrics against additive and MSD fusion algorithms*

The second experiment examines the advantage of weighing local MI differences based on block size in the quadtree. Fig. 3 illustrates the effect of weighting local MI differences by favouring low-entropy blocks $\alpha = 1$, no weighting $\alpha = 0$, and favouring high-entropy blocks $\alpha = -1$. The results show that as the minimum block size d increases, the effect of the three weighting functions tend to be equal. On the other hand, decreasing d consistently favours high-entropy weighting function.

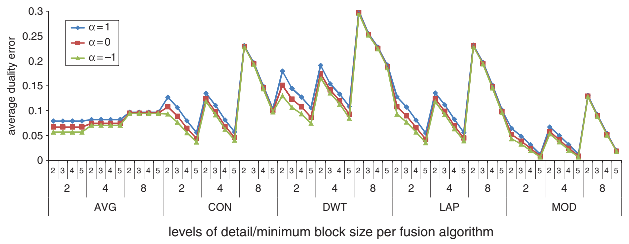

Fig. 3 *Average duality errors of three variations of* $I^{Q}_{w,\alpha|\alpha=\{-1,0,1\}}$

Conclusions: This Letter highlights the dispute created by MI's questionable results in [1] and [2]. The argument presented in this Letter explains the disputable result and suggests applying MI on smaller blocks to capture the structure similarity between source and fuse images. Using Hossny and Nahavandi's duality index [11], the experiments and results presented show that LMI records lower duality errors while following the trends of gradient-based metrics.

Acknowledgment: This research was fully supported by the Centre for Intelligent Systems Research (CISR) at Deakin University.

30 June 2010
doi: 10.1049/el.2010.1778
One or more of the Figures in this Letter are available in colour online.

M. Hossny, S. Nahavandi, D. Creighton and A. Bhatti (*Centre for Intelligent Systems Research, Deakin University, Australia*)

E-mail: mhossny@deakin.edu.au

References

1 Qu, G., Zhang, D., and Yan, P.: 'Information measure for performance of image fusion', *Electron. Lett.*, 2002, **38**, pp. 313–315
2 Cvejic, N., Canagarajah, C., and Bull, D.: 'Image fusion metric based on mutual information and Tsallis entropy', *Electron. Lett.*, 2006, **42**, pp. 626–627
3 Xydeas, C., and Petrovic, V.: 'Objective image fusion performance measure', *Electron. Lett.*, 2000, **36**, pp. 308–309
4 Piella, G., and Heijmans, H.: 'A new quality metric for image fusion'. Proc. IEEE Int. Conf. on Image Processing, (ICIP), Barcelona, Catalonia, Spain, 2003, pp. 137–176
5 Kvalseth, T.O.: 'Entropy and correlation: some comments', *IEEE Trans. Syst. Man Cybern.*, 1987, **17**, pp. 517–519
6 Tsallis, C.: 'Possible generalization of Boltzmann-Gibbs statistics', *J. Stat. Phys.*, 1988, **52**, pp. 479–487
7 Hossny, M., Nahavandi, S., and Creighton, D.: 'Comments on information measure for performance of image fusion', *Electron. Lett.*, 2008, **44**, pp. 1066–1067
8 The online resource for research in image fusion, http://imagefusion.org
9 Hossny, M., Nahavandi, S., and Creighton, D.: 'Zero and infinity in multiscale image fusion'. Proc. IEEE Int. Conf. on Image Processing, (ICIP), Cairo, Egypt, 2009, pp. 2181–2184
10 Li, H., Manjunath, B.S., and Mitra, S.K.: 'Multisensor image fusion using the wavelet transform', *Graph. Models Image Process.*, 1995, **57**, pp. 235–245
11 Hossny, M., and Nahavandi, S.: 'Image fusion algorithms and metrics duality index'. Proc. IEEE Int. Conf. on Image Processing (ICIP), Cairo, Egypt, 2009, pp. 2193–2196