Original Research Paper

SVD-Structural Similarity in the Wavelet-Gabor Domain: Improved Confidence for Face Recognition under Noise, Blur and Haze

Ohood Fadil Ismael and Zahir M. Hussain

Faculty of Computer Science and Mathematics, University of Kufa, Najaf, Iraq

Abstract: In this work we propose and investigate the performance of a new similarity measure based on Singular Value Decomposition (SVD) and structural similarity in the wavelet and Gabor domains. The reasoning behind this combination is to utilize SVD in getting independent components, wavelet decomposition to get the complex frequency features and Gabor filtering to get textural features. A comparison has been made versus correlative and structural similarity measures like SSIM (Structural Similarity Index Measure), Complex-Wavelet SSIM (CWSSIM) and FSIM (Feature-Based Similarity). In these tests, a reference image is tested for similarity against several face images in a database under adverse conditions like noise, blur and haze. A new haze formation approach has also been proposed. Similarity level and similarity confidence are taken as the performance measures. Two confidence measures, different in strength of confidence, have been proposed and tested versus a recently-proposed confidence measure that relies on the difference between the maximal similarity (best match in the database) minus the second maximum similarity (second-best match in the database). Simulation using AT&T database has shown that the proposed SVD-Structural Similarity in Wavelet-Gabor Domain (SVWG) outperforms existing measures by far. SVWG can give more robust decisions (near-optimal confidence); also, can work under more adverse conditions (lower SNR, more blur or haze) where other similarity measures fail.

Keywords: Singular Value Decomposition, SVD, Structural Similarity, SSIM, Gabor, Wavelet, Haze, Face Recognition

Introduction

Signal similarity is the science of finding similarity between signals that are probably delayed or affected by noise, blur, or other kinds of distortion (Wang *et al*., 2004; Chalom *et al*., 2013; Shnain *et al*., 2017; Aljanabi *et al*., 2018; Hassan *et al*., 2014; Hashim and Hussain, 2014). This topic has many applications in security, object recognition, medical diagnosis and many other applications.

One of the main applications of image similarity is in face recognition (Shnain *et al*., 2017; Aljanabi *et al*., 2018; Hassan *et al*., 2014; Hashim and Hussain, 2014). One of the main streams in face recognition is to recognize a given face image in the sense of similarity with some image in a large face-database. This process involves a lot of unresolved difficulties (Hashim and Hussain, 2014; Singh and Prasad, 2018; Sang *et al*., 2016;

Mahto and Yadav, 2014; Jafri and Arabnia, 2009; Seyed and Hussain, 2009).

Similarity can be 1-dimensional or 2-dimensional. Similarity measure is a mathematical technique that can detect similarity between signals in a way similar to human recognition (Wang *et al*., 2004). Similarity measures can be based on correlative, transformational or statistical analysis (Shnain *et al*., 2017; Aljanabi *et al*., 2018). The conventional metrics such as the Peak Signal-To-Noise Ratio (PSNR) and the Mean Squared Error (MSE) operate directly on the intensity of the image and they do not correlate well with the subjective fidelity ratings. Thus, many efforts have been made to design Image Quality Assessment (IQA) metrics that are based on the functions performed by The Human Visual System (HVS). Such kinds of models emphasize the importance of HVS sensitivity to different visual

signals, such as luminance, contrast and frequency content (Wang *et al*., 2004; Chalom *et al*., 2013; Shnain *et al*., 2017; Aljanabi *et al*., 2018).

Since 2004, several improvements have been proposed to SSIM. One of the most important improvements was the Complex Wavelet Structural Similarity, CWSSIM (Sampat *et al*., 2009), which was defined in a way similar to SSIM, but instead of pixels, they used discrete wavelet coefficients. In 2011, Zhang et al. proposed a feature-based image similarity index measure (FSIM), which utilized phase congruency and gradient features (Zhang *et al*., 2011).

In 2017, A novel feature-based measure (FSM) that combines the best features of the well-known SSIM and FSIM approaches was proposed, striking a balance between their performances for similar and dissimilar images (Shnain *et al*., 2017). In 2018, an information-theoretic measure based on the joint histogram was proposed in (Aljanabi *et al*., 2018).

In another direction, independent component analysis (via singular value decomposition, SVD) played a significant role in face recognition (Turk and Pentland, 1991; Slavković and Jevtić, 2012; Çarıkçı and Özen, 2012; Zhang , 2011). It decomposes the face matrix into statically independent components that would constitute a basis for the face images, then all images are projected onto these basis functions and minimal distance is searched for recognition.

On the other hand, transformational approaches for face recognition and finding image features proved to be efficient (Lajevardi and Hussain, 2009c; Hussain and Boashash, 2001; Sampat *et al*., 2009; Zhang *et al*., 2011; Turk and Pentland, 1991; Slavković and Jevtić, 2012; Çarıkçı and Özen, 2012; Zhang, 2011; Gabor, 1946). One of these approaches is to use Gabor filtering. Gabor filter has been proposed in 1946 by Gabor (Gabor, 1946) and extended into two-dimensional function by Daugman (Daugman, 1985). Gabor filter is resistant against a moderate change of illumination; also, it is shown that Gabor filter can detect textural features in the image (Gabor, 1946; Lajevardi and Hussain, 2009a; He *et al*., 2011).

Shnain *et al*. (2017; Aljanabi *et al*., 2018), a measure for confidence in the recognition capability of similarity measure was introduced. It is the difference between the maximum similarity (which should be at the database face image that matches the reference face) and the second-best match. This difference, if small, could cause confusion in the recognition process.

In this work, we test existing similarity measures under measures of confidence. It is found that most of the existing measures can be highly affected by pose, noise, blur and haze. Hence, we propose a new similarity measure with near-optimal performance regarding confidence. The new measure extracts statistical, structural, textural and complex-frequency features. The proposed measure, called SVD-Wavelet-Gabor measure (SVWG), outperforms all existing similarity measures in the face recognition process, with highest confidence. In addition, we designed two new measures for recognition confidence.

This paper is organized as follows. Section 2 surveys related works with a brief background. Section 3 presents a new similarity measure based on SVD and SSIM as applied to wavelet and Gabor domains. Section 4 explores the concept of recognition confidence and presents two measures for confidence. Section 5 presents results of testing the proposed SVWG versus existing measures. Test environment includes change of pose, noise, blur and haze, where a new approach for the formation of haze is presented. Some future directions are also discussed.

Background and Related Work

Automatic detection of similarity between images plays an important role in many image processing applications such as image enhancement, object recognition and identity recognition.

One of the major challenges in measuring similarity for face recognition is that test image may have low resolution or distortion due to changes in illumination or noise. Other challenges are facial expressions and different head poses. Official security systems (e.g., in airports) do not rely completely on face recognition systems since these systems are still suffering from failures due to different facial expressions, noise, illumination and aging. However, more efficient tool for face recognition can be more helpful for security or recognition applications.

During the last few years, many studies on image similarity have been presented to design a similarity measure that can achieve high success rates when applied in image recognition. An important image similarity measure was designed by Wang and Bovik in 2004 (Wang *et al*., 2004), called structural similarity index measure (SSIM). This measure has been designed using statistical relations between the two images. However, SSIM can give reasonable amount of similarity between two different images, a fact that causes confusion in the recognition process (Shnain *et al*., 2017; Aljanabi *et al*., 2018).

Different attempts for improving SSIM have been proposed, including the Complex Wavelet Structural Similarity, CWSSIM (Hussainand Boashash, 2001), which used discrete wavelet coefficients instead of pixels in the similarity measure. In many research directions, wavelet transforms proved useful in extracting spectral components of images and signals in the complex frequency domain.

In 2011, an improved version of SSIM was proposed by (Zhang *et al*., 2011). This measure was called feature similarity index measure (FSIM), which used low level

features (in accord with HVS), specifically phase congruency and gradient magnitude. Phase congruency approach extracts features at points where Fourier transform of the signal has maximal phase, while gradient is measures by 2D-convolution with specific masks like Sobel or Prewitt operators. However, FSIM can give confusing results in finding the similarity between two different images.

Another direction in feature analysis has used independent component analysis (via singular value decomposition, SVD) to extract features based on statistically - independent components of the image (or, signal). This approach played significant role in pattern recognition (Turk and Pentland, 1991; Slavković and Jevtić, 2012; Çarıkçı and Özen, 2012; Zhang, 2011).

Other approaches for image similarity gave good results, including rational functions of existing measures and information-theoretic measures (Shnain *et al*., 2017; Aljanabi *et al*., 2018).

In general, several similarity metrics has been proposed over the last decade and applied in pattern recognition; however, each one has its own strengths and drawbacks. The common problem in image similarity for face or pattern recognition is the non-trivial similarity between the reference image and some of the unrelated (different) database images, causing confusion (less confidence) in recognition.

In this paper, we proceed from four different directions of feature extraction, where we design a high-performance measure based on statistical, structural, textural and complex-frequency features. The new measure utilizes statistically-independent components (by applying SVD), then extracts structural components in the Gabor and wavelet domains using SSIM approach, noting that Gabor domain contains textural features, while the wavelet domain reveals spectral properties. The versatileness of feature sources and methods gives near-optimal strength to this measure. The proposed measure reduces confusing similarities between unrelated images. The proposed measure has been tested versus some existing ones using new confidence measures under various adversities like pose, noise, blur and illumination change (haze).

The sub-sections below present a brief background on some important existing methods of similarity that are of relevance to this research.

Structural Approaches

2D-Correlation and Structural Similarity Index Measure

This important similarity measure was introduced in 2004 by Wang and Bovik (Wang *et al*., 2004), SSIM considered image similarity features as a combination of three factors: correlation, luminance and contrast. This measure proved successful in image quality assessment

and in many other image processing algorithms. SSIM used image statistical properties to find a distance measure between two images. Specifically, it utilizes mean, variance and co-variance between a reference image and another image as follows:

$$\rho(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{1}$$

where, $\rho(x, y)$ is the similarity value between a reference image x and the image under test y; where y is usually another version of x (probably affected by distortion); and μ_x, μ_y, σ_x and σ_y are the means and variances of the relevant images x and y. The constants c_1 and c_2 are small numbers. The quantity σ_{xy} represents the co-variance between the pixels in images x and y. In our opinion, SSIM is a generalization to the 2D correlation $\Re(x, y)$ between images x and y given by (Lau and Hussain, 2005):

$$\Re(x, y) = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} = \frac{\sum_{k=1}^{N}(x_k - \mu_x)(y_k - \mu_y)}{\sqrt{\sum_{k=1}^{N}(x_k - \mu_x)^2} \cdot \sqrt{\sum_{k=1}^{N}(y_k - \mu_y)^2}} \tag{2}$$

FSIM: Feature Similarity Index Measure

The Feature Similarity Index Measure (FSIM) was proposed in (Zhang *et al*., 2011). It is based on the fact that human visual system (HVS) recognizes images based on their low-level features. The main local feature of FSIM is the Phase Congruency (PC), while the second important feature is the Gradient Magnitude (GM). It is found by experiments that HVS recognizes that good (discernable) features can be extracted from points (pixels) corresponding to frequencies (obtained after taking 2D-Fourier transform) that have similar (congruent) phases. For example, frequency components of edge-like structures in the image are most likely to have same phase. However, since PC is contrast-invariant, the second feature (GM) is considered to compensate for contrast variation. The FSIM index between two images x and y is defined as follows:

$$\wr\{x, y\} = \frac{\sum_{(i,j)} S(i, j) \cdot PC(i, j)}{\sum_{(i,j)} PC_m(i, j)} \tag{3}$$

where, (i, j) represents the location of the pixel $x(i, j)$, PC is the phase congruency and S is the overall similarity after combining the phase congruency similarity $SPC(i, j)$ with the gradient similarity.

Transformational Approaches

Time-Frequency Distributions (TFDs) and the Wavelet Transform (WT)

Fourier Transform (FT) of a time signal $x(t)$ reveals the frequency content of the signal by decomposing the signal using complex sinusoids as follows:

$$X(f) = F\{x(t)\} = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt \qquad (4)$$

However, *FT* cannot reveal the time information associated with a specific frequency. This drawback enhanced research in the time-frequency domain (Hussain and Boashash, 2001). One of the most important Time Frequency Distributions (TFD's) is the Wavelet Transform (WT), which is a time-frequency representation of signals. While not all TFD's are invertible, a big advantage of WT over many other TFD's is invertibility. WT proved to be successful in revealing spectral features of signals. Instead of sustainable waves like sinusoidal waves as in the case of Fourier Transform, WT is based on decomposing signals using decaying waves (small waves, or wavelets), all are shifted and dilated versions of a specific wavelet called mother wavelet. The Continuous Wavelet Transform (CWT) of a signal $x(t)$ using a mother wavelet $\psi(t)$ is given by:

$$W_x^{\psi}(t,s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(\lambda) \cdot \psi * \left(\frac{\lambda - t}{s} \right) d\lambda \qquad (5)$$

where, λ is a representation of time inside the convolution integral, $\psi*$ is the complex conjugate of the wavelet ψ and $s \in \mathbb{R}^+ = \mathbb{R} - \{0\}$ is called the "scale", which we expect to be inversely related to the radian frequency $\omega = 2\pi f$.

The above continuous wavelet transform can be discretized to give the Discrete Wavelet Transform (DWT), which can be implemented (as 1D DWT) by passing the signal $x(t)$ through a low-pass filter followed by down-sampling with a factor of 2 (giving approximation coefficients) and a high-pass filter then down-sampling by a factor of 2 (giving detail coefficients). These filters differ according to the analyzing wavelet (The MathWorks, 2019).

The 2D DWT (for images) can be designed based on 1D DWT via tensor products and it results into a decomposition of approximation coefficients at level k into four components: low-pass component that contains the approximation coefficients at level $k +1$ and three high-pass components that contain the detail coefficients in three directions (horizontal, vertical and diagonal). Note that approximation at level $k = 0$ is equivalent to the original image (The MathWorks, 2019).

CWSSIM: Complex Wavelet SSIM

This similarity measure was introduced in 2009 in an attempt to design measure that are compatible with HVS (Sampat *et al*., 2009). Instead of using pixel values to find structural properties, CWSSIM uses wavelet coefficients of the two images **x** and **y**. The analyzing wavelets are symmetric complex wavelets generated from a mother wavelet of the form:

$$\psi(t) = f(t)e^{j2\pi f_o t} \qquad (6)$$

where, $j = \sqrt{-1}, f(t)$ is a symmetric low-pass filtering function, modulated by a complex sinusoid of frequency f_o. If the wavelet coefficients of the two images x and y are given by the two sets:

$$C_x = \left\{ c_{xk} \big| k = 1 : N \right\} and \, C_y = \left\{ c_{yk} \big| k = 1 : N \right\}$$

where, corresponding coefficients in the above two sets are related to the same wavelet sub-bands in the wavelet domains of the two images under test, then CWSSIM index is given by:

$$W(\pmb{x}, \pmb{y}) = \frac{2\gamma_{xy} + k}{\gamma_x + \gamma_y k} \qquad (7)$$

where:

$$\gamma_{xy} = \left| \sum_{k=1}^{N} c_{xk} \cdot c_{yk}^* \right|$$

$$\gamma_x = \sum_{k=1}^{N} |c_{xk}|^2 ; \gamma_y = \sum_{k=1}^{N} |c_{yk}|^2$$

with k a small positive number, inserted for stability.

Gabor Filtering

Gabor filter has been widely used in image processing, where filtering an image $\pmb{x} = \{x (k, l)\}$ with a Gabor filter $\pmb{g} = \{g (k, l)\}$ (via 2D convolution) will extract textural features. Image texture refers to repetitive spatial patterns in the pixels of the image, hence it cannot be defined for a single point (Lajevardi and Hussain, 2009a). The 2D impulse response of Gabor filter can be represented as follows (Lajevardi and Hussain, 2009a):

$$g(k,l) = \frac{1}{2\pi\sigma_k\sigma_l} \exp\left[-\frac{1}{2}\left(\frac{\tilde{k}^2}{\sigma_k^2} + \frac{\tilde{l}^2}{\sigma_l^2} \right) \right] \times \exp\left[jw\tilde{k} \right] \qquad (8)$$

where:

$$\tilde{k} = k \cdot \cos(\theta) + l \cdot \sin(\theta)$$
$$\tilde{l} = -k \cdot \sin(\theta) + l \cdot \cos(\theta)$$

Noting that (k, l) is the pixel location, $w = 2\pi F$ is the of radian frequency of the filter wave in the direction of the rotated axis $\tilde{k}$ (where the frequency F is the reciprocal of the wavelength T), the imaginary root is $j = \sqrt{-1}$, θ is the orientation of the filter, σ_k^2 and σ_l^2 are the variances of the filter in the k and l directions, which decide the spread of the filter kernel in the two spatial directions. Now the Gabor-filtered image would be given by:

$$x_g = x \circledast g \cdot \tag{9}$$

ICA-Based Approaches

SVD Similarity Measure

This measure is obtained by using Singular Value Decomposition (SVD) to decompose the matrix whose columns are the 1D versions of database images to calculate a basis of independent vectors for this 1D database (Turk and Pentland, 1991; Slavković and Jevtić, 2012; Çarıkçı and Özen, 2012). Then all images (including the reference image) are projected onto this set of basis vectors (coordinates) and a distance measure (normally Euclidean) is defined between each image of the database (expressed in the new coordinates) and the reference image (expressed in the new coordinates), where recognition is considered at the minimum distance.

Let the database of face images be $S = \{x_k | k = 1 : P\}$, including P persons and the reference image be y, all of the same size $m \times n = N$ pixels. The SVD implies that all images are converted to 1D columns and inserted in one $N \times P$ global matrix as follows:

$$C = \begin{bmatrix} C_1 C_2 ... C_P \end{bmatrix}$$

while the 1D column that represents the reference image will be denoted by R. As the mean does not convey any information but it may confuse calculations by producing large numbers, the column mean is subtracted to get the modified global matrix s as follows:

$$c = \begin{bmatrix} c_1 c_2 \cdots c_P \end{bmatrix}$$

where, $c_k = c_k - \mu$, with $\mu = \dfrac{\sum_{k=1}^{P} c_k}{P} =$ is the average column vector. The 1D reference image will be given by $r = R - \mu$. Now the matrix c (with size $N \times P$) is decomposed via SVD as follows:

$$c = U_{N \times N} S_{N \times P} V_{P \times P}^T \tag{10}$$

where, U and V are orthonormal matrices (i.e., $UU^T = I$, $VV^I = I$). Being orthonormal, this means that the columns of

U are independent, hence, we can choose any (e.g., the first) P of these columns to build a basis $B_{N \times P} = [u_1 \; u_2 \; \ddot{} u_P]$ for the P database images. If we project the database images $c_{N \times P}$ onto the new basis we get the new coordinates of the P database images in the P- dimensional vector space:

$$Z_{P \times P} = \left(B_{N \times P}\right)^T * c_{N \times P} = \begin{bmatrix} z_1 z_2 \cdots z_P \end{bmatrix} \tag{11}$$

while the reference (test) image will be projected as:

$$w_{P \times 1} = \left(B_{N \times P}\right)^T * r_{N \times 1} \tag{12}$$

Now the Euclidean distance between the reference image w and each of the database images $\{z_i | i = 1 : P\}$ in the new coordinates will be:

$$d_i = \| w - z_i \| \tag{13}$$

SVD-Similarity is based on distance, normalized to maximum distance $d_o = \max \{d_i | i = 1 : P\}$ as follows:

$$\partial(x_i, y) = 1 - \frac{d_i}{d_o} \tag{14}$$

and recognition is based on the minimum distance (hence, maximal similarity). However, this approach suffers from the drawback that not all SVD components (matrices) are utilized in the recognition process, causing some information to be lost.

SVD-Structural Similarity in Wavelet-Gabor Domain (SVWG)

In the review above, four different directions of feature extraction are utilized in designing well known similarity measures: statistical, structural, textural and complex-frequency features. This fact inspired the design of a high performance measure based on a specific combination of these features. This measure applies SVD analysis to extract similarity between statistically-independent components, which is combined with structural similarities in the Gabor and wavelet domains using SSIM, hence the name SVWG. Combining features from unrelated sources and using unrelated methods gives almost optimal performance in face recognition. The proposed measure outperforms existing measures in annihilating the unwanted similarities between different images, giving more confidence for the recognition process.

The proposed measure is designed as follows.

Input: A reference image y and a database of P images $S = \{x_k | k = 1 : P\}$. SVD-Similarities:

$$\{\partial(x_k, y) | k = 1 : P\}$$

Gabor-Filtered images:

$$y_g = y \otimes g; \tag{15}$$

$$\{x_{g,k} = x_k \otimes g \,|\, k = 1:P\} \tag{16}$$

Gabor-filtered images in magnitude and phase:

$$y_g = y \otimes g \rightarrow A(y_g), \varphi(y_g) \tag{17}$$

$$\{x_{g,k} = x_k \otimes g \,|\, k = 1:P\} \rightarrow$$
$$\{A(x_{g,k}), \varphi(x_{g,k}) = |k = 1:P\} \tag{18}$$

SSIM in Gabor domain:

$$\{\rho_{ga}(x_k, y) = \rho(A(x_{g,k}), A(x_g)) \,|\, k = 1:P\} \tag{19}$$

$$\{\rho_{gp}(x_k, y) = \rho(\varphi(x_{g,k}), \varphi(x_g)) \,|\, k = 1:P\} \tag{20}$$

$$\{\rho_g(x_k, y) = \rho_{ga}(x_k, y) \cdot \rho_{gp}(x_k, y) \,|\, k = 1:P\} \tag{21}$$

wavelet analysis of images; horizontal, vertical and diagonal (ψ is Haar or Daubechies 1):

$$W_{x_k}^{\psi}(t,s) = [Ay, Hy, Vy, Dy] \tag{22}$$

$$\{W_{x_k}^{\psi}(t,s)\} = \{[Ax_k, Hx_k, Vx_k, Dx_k] \,|\, k = 1:P\} \tag{23}$$

SSIM in wavelet domain:

$$\begin{cases} \rho^{\psi}(x_k, y) = \rho(Ax_k, Ay) \cdot \rho(Hx_k, Hy) \\ \cdot \rho(Vx_k, Vy) \cdot \rho(Dx_k, Dy) \,|\, k = 1:P \end{cases} \tag{24}$$

Now SVWG in the wavelet-Gabor domain is given by:

$$\begin{cases} \eta^{\psi}(x_k, y) = \partial(x_k, y) \cdot \rho_g(x_k, y) \\ \cdot \rho^{\psi}(x_k, y) \,|\, k = 1:P \end{cases}. \tag{25}$$

A flow-chart of the proposed approach is shown below.

Recognition Confidence: More Conditions and Measures

Although two similarity measures may success in recognizing a face reference image versus a database of face images, still the confidence in this decision can vary between the two measures.

Suppose two similarity measures (M_1 and M_2) correctly recognized a reference (test) face image from a database. This recognition does not mean that the two measures have the same performance. For example, if M_1 produces significant similarities between the test image and another group of N database images, then the suspected faces would be not only the one with maximal similarity, but there would be $1 + N$ of possible (confusing) faces. Now if M_2 produces significant similarity with the correct face image in the database while giving trivial similarities with the rest of the database images, then M_2 is giving the correct answer with more confidence than M_1 (which gave several confusing faces in addition to the best match).

Confidence measure was first introduced as performance measure in (Shnain *et al*., 2017) and (Aljanabi *et al*., 2018). This measure, referred to as d_M, is the difference between the maximum similarity (which, under correct recognition, should be at the database face image that best matches the reference face) and the second-best match when using a specific similarity measure M. Based on this confidence measure, the quality of recognition varies from one measure to another, where the similarity measure M which gives maximal d_M is considered the most confident.

Still, the confidence in recognition can be weakened under stronger conditions of confidence, under which only a few similarity measures can be confident of their decision. Finding the similarity using measure M between a test image and a database of N face images will produce a similarity set $\mathbb{S}_M = \{s_i \,|\, i = 1:N; 0 \le s_i \le 1\}$ that contains v similarity values, ranging between 0 and 1. Note that, without loss of generality, all measures are assumed to be normalized to the range (Wang *et al*., 2004). where the value of 1 is given to the similarity with the best match (under correct recognition, otherwise 0 is assigned under wrong decision). If the similarity level is fixed at a value k, where $0 \le k \le 1$, then a number n_k of these similarity values of $\mathbb{S}_M$ can exceed the level k, giving confusion that depends on the value of k (higher is k the less is the confidence if a peak other than the bet match exceeds k). It is clear that n_k is method-dependent. The larger n_k the weaker in confidence is the measure under test at level k.

Combining the above two directions for confidence, two measures for confidence are designed based on the first measure above d_m and the number n_k of confusing similarities at level k as follows. The symbol P refers to the total number of persons in the database.

Measure -1, m_1 (Shnain *et al*., 2017; Aljanabi *et al*., 2018):

$$m_1(M) = d_M \tag{26}$$

Measure -2, m_2:

$$m_2(M) = d_M \frac{P - n_k + 1}{P} \tag{27}$$
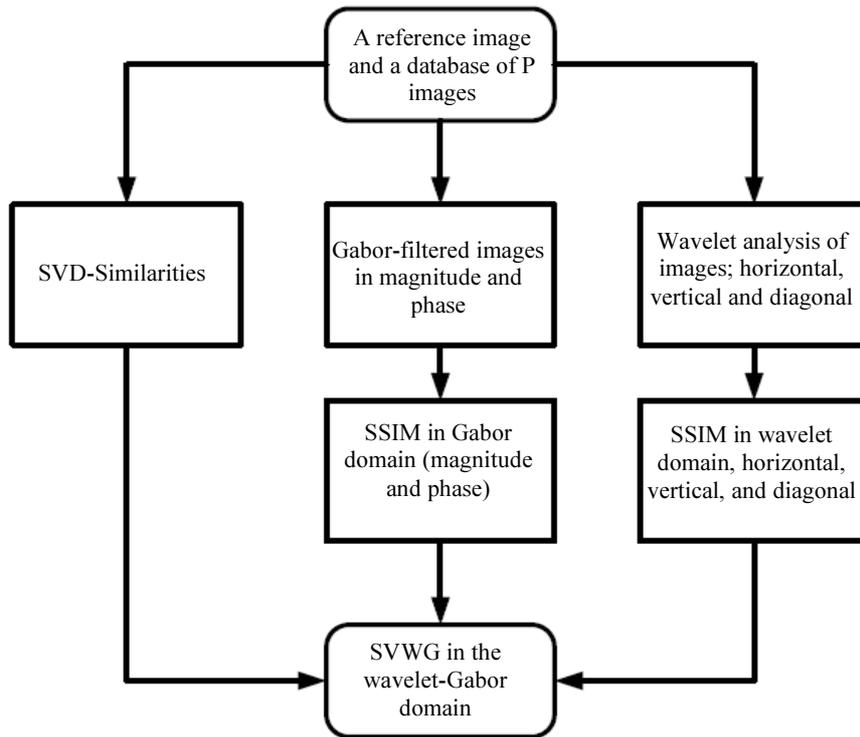
Fig. 1: A flow-chart of the proposed SVWG similarity

n_k = number of peaks above similarity level k
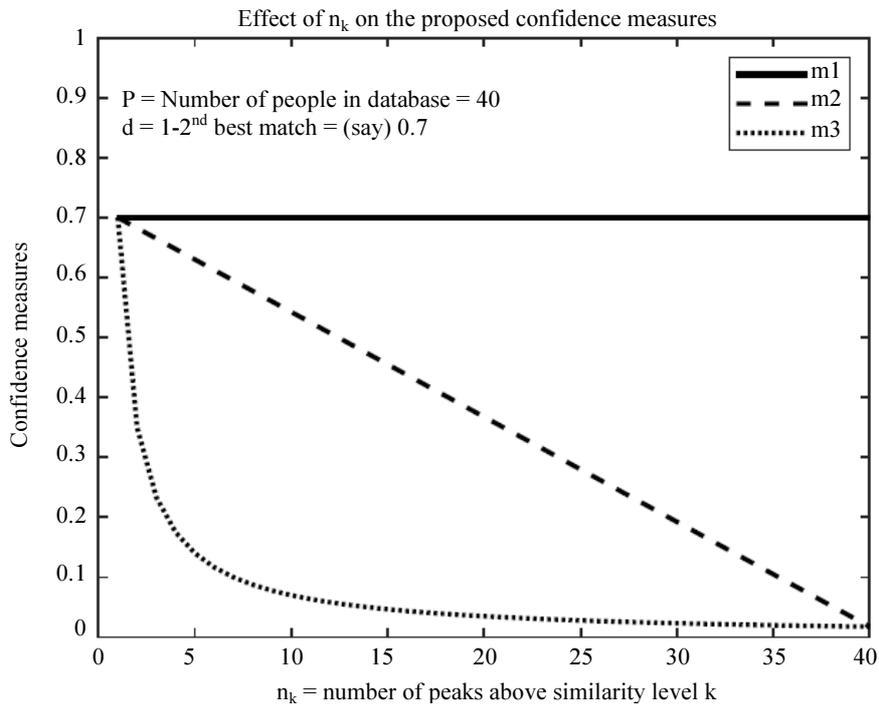
Fig. 2: Effect of nk on the proposed measures, where nk is the number of peaks above similarity level k in the recognition process. Such peaks reduce confidence in recognition. The main confidence factor, d, which is the distance between the maximum (normalized to 1) and the 2nd best match (in case of correct recognition) is fixed in this figure at 0.7. It is clear that measure-2 is stronger than measure-1, and measure-3 is stronger than all, as it is highly affected by the number of confusing peaks. Note that the lower is k, the higher nk is expected

1215

Note that, under successful recognition, when $n_k = 1$ (i.e., only the best match exceeds level k), then $m_2(M) = m_1(M)$.

Measure -3, m_3:

$$m_3(M) = \frac{d_M}{n_k} \qquad (28)$$

Again, when $n_k = 1$ (under successful recognition), then $m_3(M) = m_1(M)$.

Note that measure -2 is stronger than measure -1 = d_M, while measure-3 is the strongest as shown in Fig. 2.

Results and Discussion

The proposed measure has been tested versus some powerful existing ones using new confidence measures under various adversities like different pose, facial expression, noise, blur and illumination change (haze). Note that an image with a different pose or facial expression for the same person would be considered by any naive computer program as a different person. Hence, deeper analysis of the image and its properties are required by similarity measures for correct recognition under adverse circumstances.

This work has utilized a well-known face image database called AT&T Database of Faces (formerly "The ORL Database of Faces") (AT&T, 2019). This database includes the face image for $P = 40$ persons, each person has 10 different poses and facial expressions, hence a total of 400 face images.

Gaussian Blur

Blur is a smoothing effect on the image that reduces fine details (like shar edges); it has been simulated in this work via filtering the reference image x with a 2D Gaussian filter. Under blur, the original pixel value x_o is transformed into a weighted average of that value plus the weighted values for other pixels in the neighborhood of x_o. With a Gaussian filter, the new pixel value is composed of the original pixel's value x_o multiplied by the maximum weight, plus the neighboring pixels which are multiplied by smaller weights that decay as their distance to the original pixel increases.

If the 2D Gaussian filter has the impulse response $G = \{G(i, j)\}$, then the blurred version x_b of the original image $x = \{x(i, j)\}$, is given by the 2D-convolution:

$$x_b = x \circledast G. \qquad (29)$$

Haze Formation

Haze is a phenomenon caused by interaction between the light (coming from objects to the camera) and atmosphere; this effect results in a change of illumination

and contrast (He *et al*., 2011; Park *et al*., 2014; Lee *et al*., 2016). Most importantly, ambient light could be reflected into the line of sight by atmospheric particles. The model that is widely used in this respect is given by (He *et al*., 2011):

$$y(i, j) = x(i, j)t(i, j) + A[-t(i, j)] \qquad (30)$$

where, y is the observed intensity (haze-polluted image intensity that reaches the camera), x is the original scene illumination (original haze-free image), A is the atmospheric light and t is the atmosphere transmission map. Note that (i, j) are the pixel coordinates. When the atmosphere is homogenous, the transmission map t can be expressed as:

$$t(i, j) = e^{-\beta d(i,j)} \qquad (31)$$

where, β is the scattering coefficient of the atmosphere and d is the scene depth (distance from the camera). This equation indicates that the scene radiance is attenuated exponentially with the depth.

The above model is complicated and requires real data for processing, especially the transmission map. In this work we consider only one aspect of haze, which is a change of illumination by extra atmospheric light, pushing the image illumination towards white. Hence, we propose a special formula for haze effect which is a non-linear transformation (distortion) on the original pixel value x; a transformation that incorporates additive and multiplicate properties of haze as follows:

$$f(x) = \frac{x + k}{1 + k} \qquad (32)$$

assuming that x is in the range [0,1] (Wang). The value k is called the haze level. Figure 3 shows the effect of the proposed haze formula on pixel values. Note that when $k \to \infty$; then $f(x) \to 1$.

Simulation Results and Discussion

The proposed SVWG measure outperforms existing measures by far in giving less confusion while recognizing the best match for a given face image. In addition to confidence, the proposed measure (SVWG) gives a reliable similarity between any two similar images even under adverse conditions like noise, blur and haze. In other words, SVWG produces maximal similarity when the images are similar (even under adverse conditions like pose and noise), while giving near-zero similarity when the images are dissimilar. Such properties are highly needed in security applications while checking the identity of a specific face image in a big database.

In the results below for the recognition process, a different pose is considered for the reference image. This different pose is subjected further to noise, blur and haze.
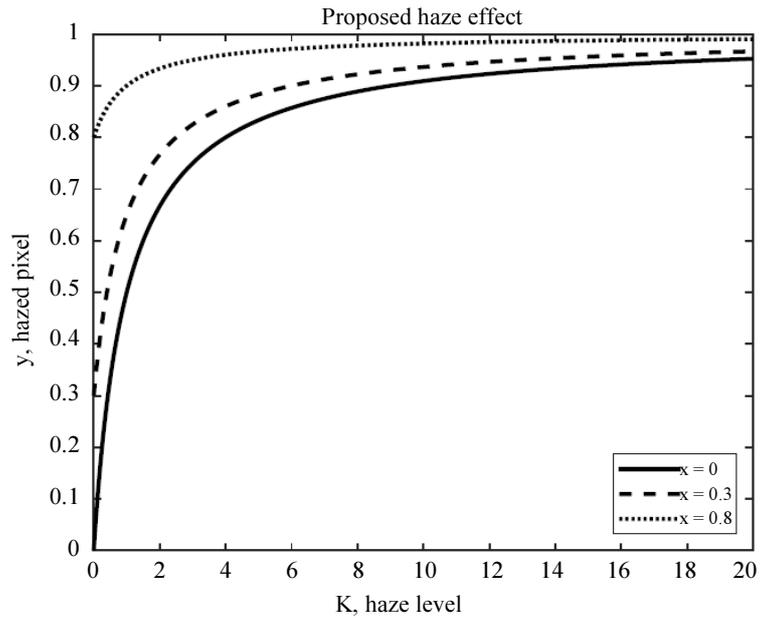
Fig. 3: Proposed haze effect as represented by the non-linear transformation f(x)=(x+k)/(1+k). Note that as haze level k increases, f(x) approaches 1 (white).
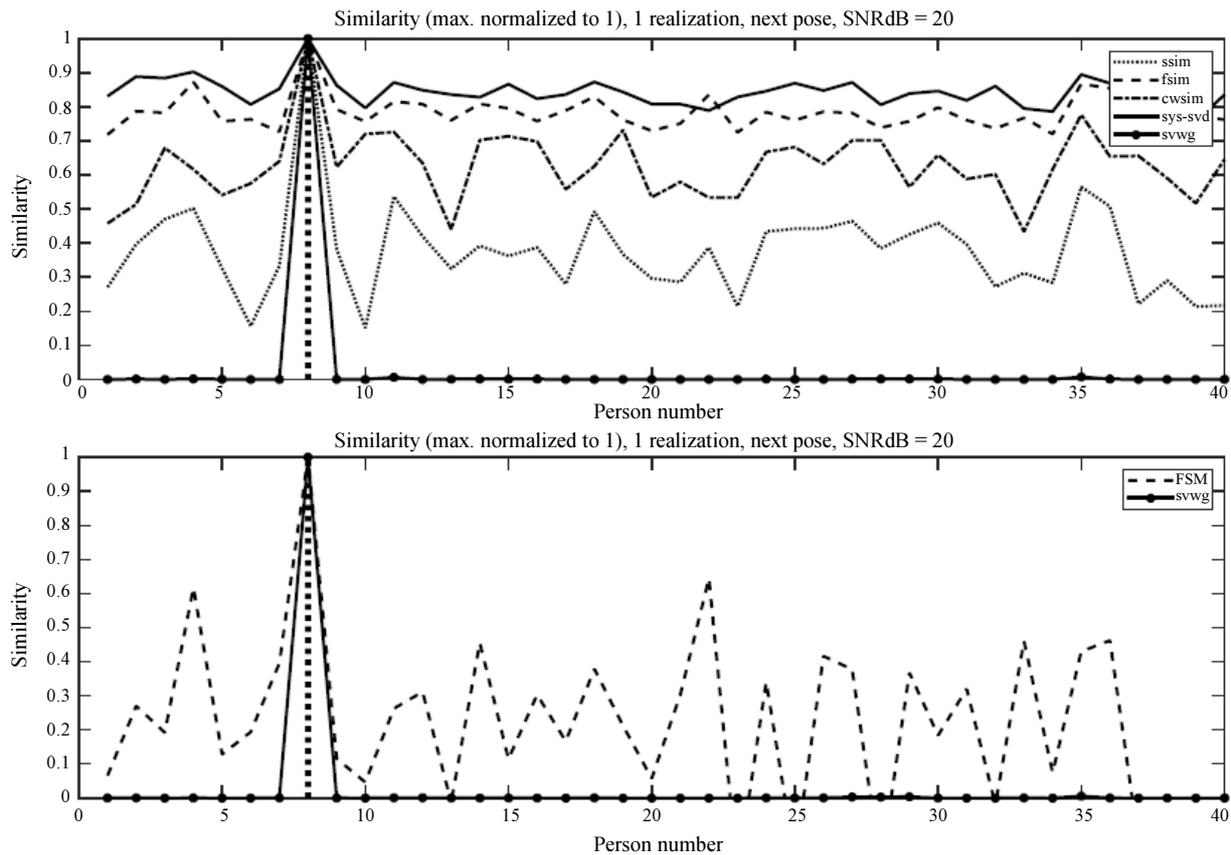
Fig. 4: Face recognition under noise using similarity. Reference is 8th person in AT&T Database. Note that a different pose is used for testing similarity measures. Above: Recognition based on similarity measures, with SNR at 20 dB. Below: A comparison with the recently proposed FSM [3], at SNR of 20dB

1217
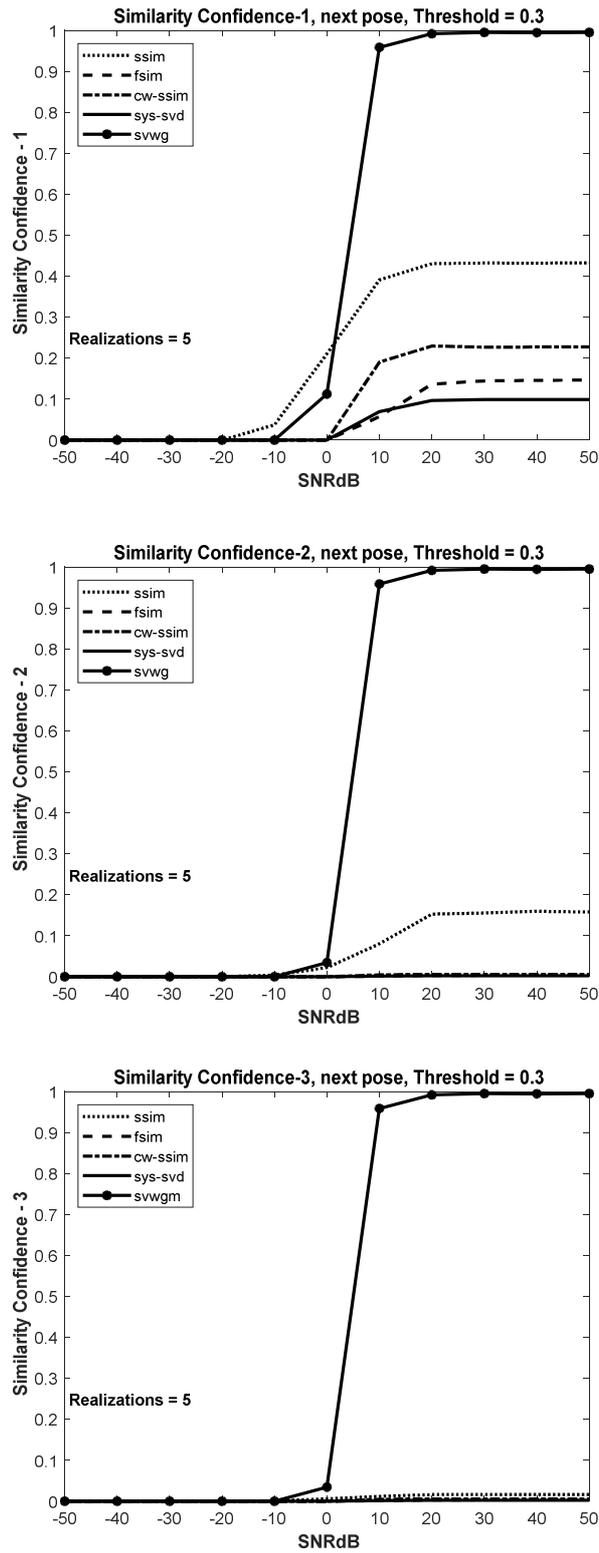
Fig. 5: Confidence in face recognition under noise using similarity threshold of 0.3. Three confidence measures have been used. Note that a different pose is used for testing recognition confidence, as in Figure 4. Left, middle, and right figures represent the three confidence measures as shown. Note that the 3rd measure is stronger than the 2nd, which is in turn stronger than the first. Only the proposed svwg survives under the 3rd. Five realizations are performed

1218
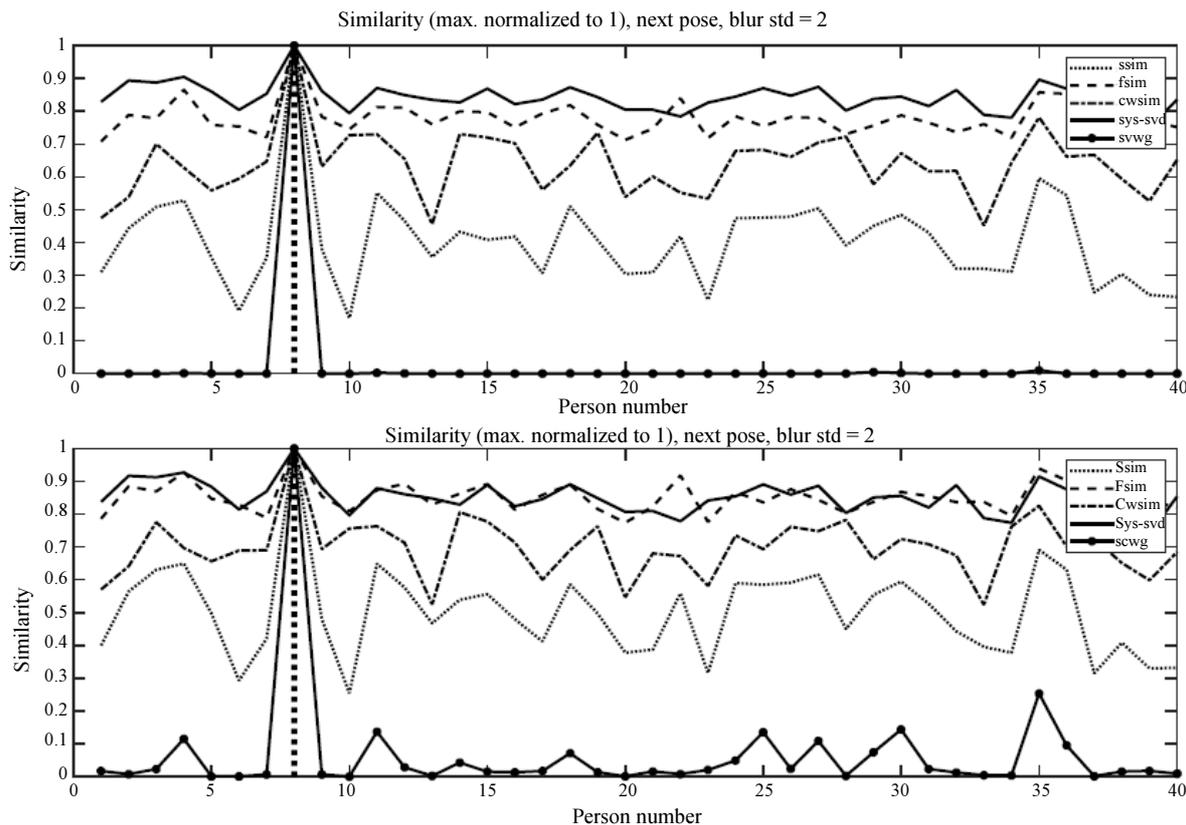
Fig. 6: Face recognition under blur (no noise) using similarity. Reference is 8th person in AT and T Database. Note that a different pose is used for testing similarity measures. Recognition is based on similarity measures, with two values for blur filter STD

1219
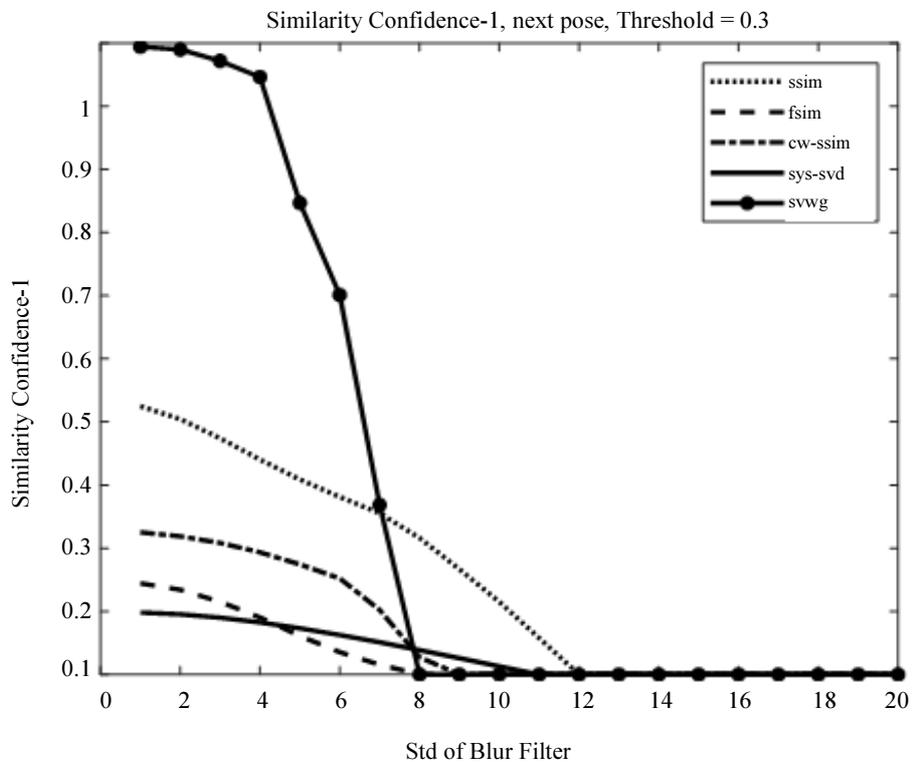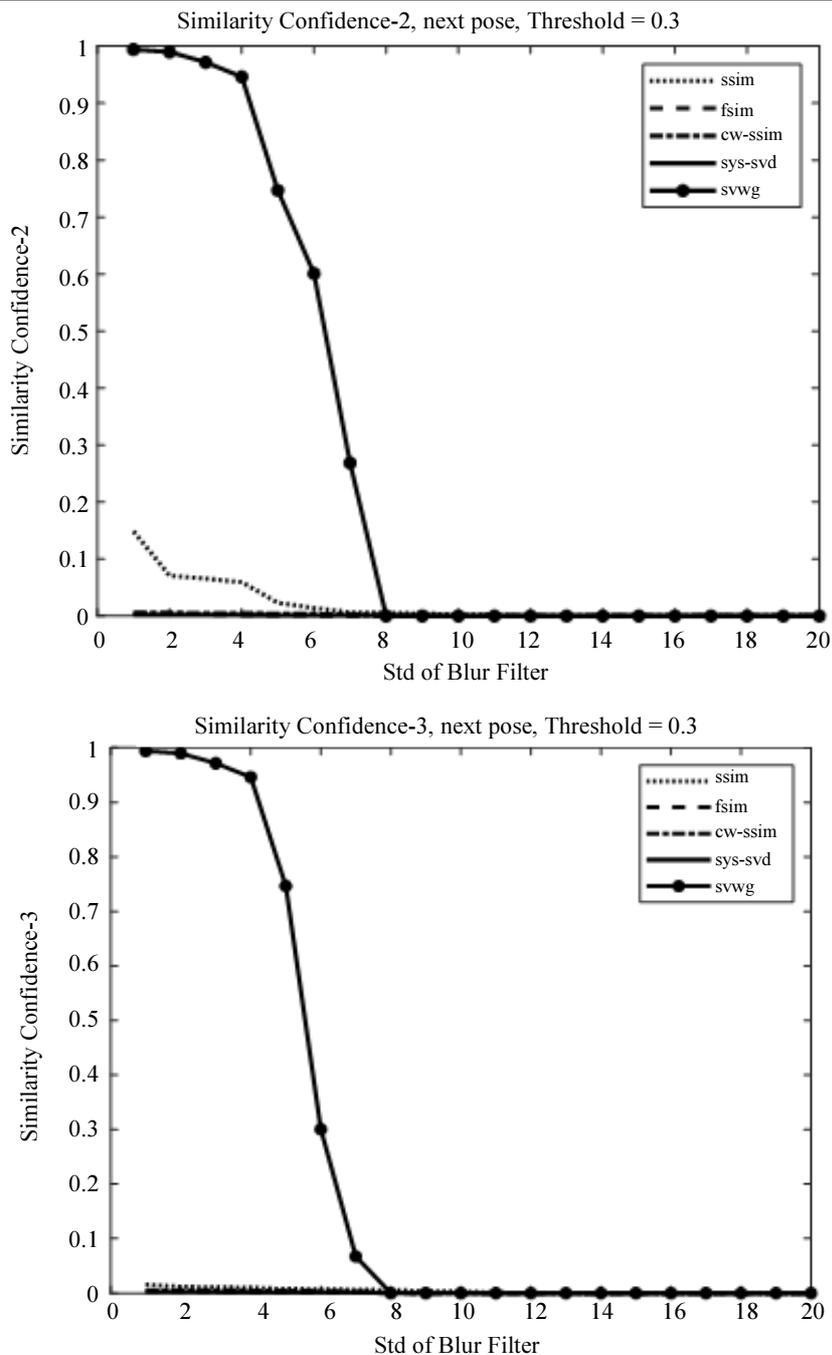
Fig. 7: Confidence in face recognition under blur using similarity threshold of 0.3. Three confidence measures have been used. A different pose is used for testing recognition confidence, as in Figure 6. Left, middle, and right figures represent the three confidence measures versus blur filter SDT as shown. Note that the 3rd measure is stronger than the 2nd, which is in turn stronger than the first. Only the proposed svwg survives under the 3rd

Note that a different pose may include different facial expression (e.g., a smile).

It is found that the proposed SVWG works better Haar and Daubechies wavelets, while we used the function "imgaborfilt" for Gabor filtering with wavelength = 2 and angle of 92 degrees. Figure 4 shows the face recognition process using AT&T face database (AT&T) under noise using different similarity measures under two values for PSNR. A different pose is considered. We note that SVWG gives almost zero similarity for images related to people other than the targeted person (reference person), while other similarity measures give a non-trivial amount of similarity with wrong people. Although all agree on the best match, but with different levels of confidence.

1220
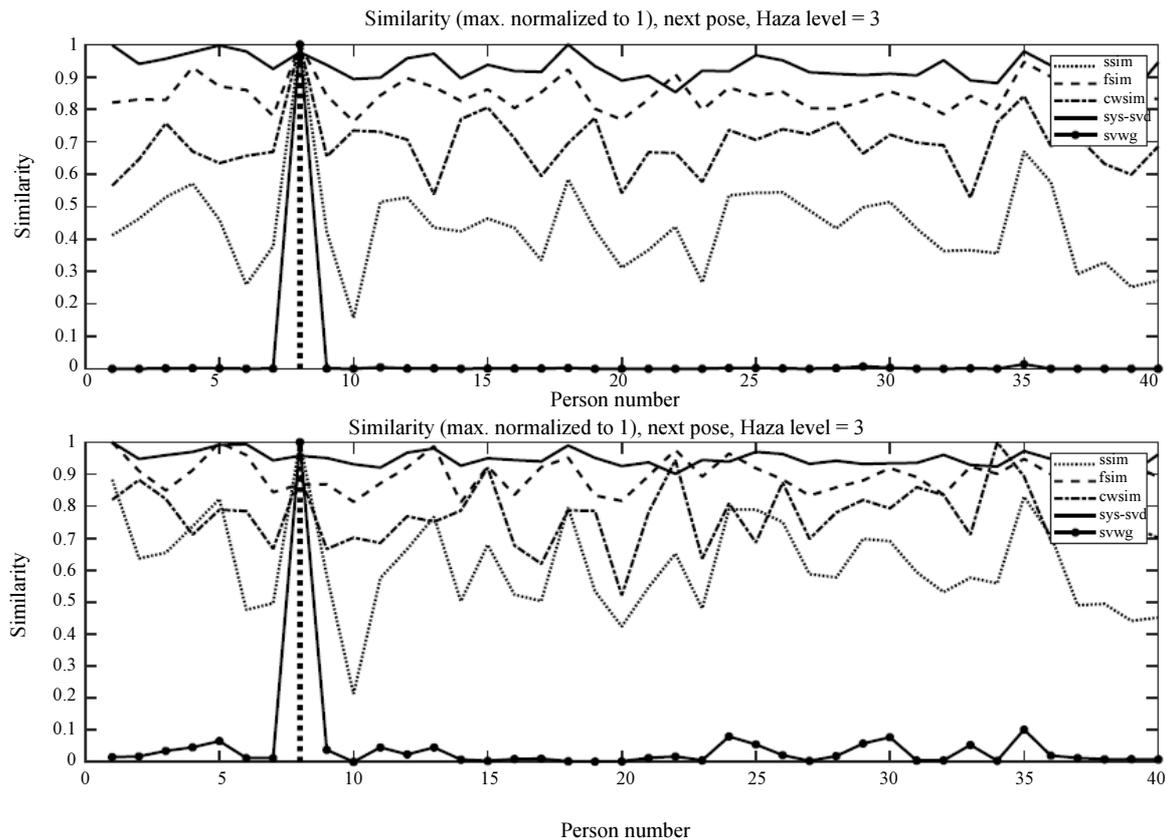
Fig. 8: Face recognition under haze (no noise) using similarity. Reference is the 8th person in AT and T Database. Note that a different pose is used for testing similarity measures. Recognition is based on similarity measures, with two values for haze level. Note that at higher levels of haze, all similarity measures might give higher similarity with non-similar images (or even fail, except for SSIM and SVWG), hence reducing recognition confidence

Note that Fig. 4 also includes a performance comparison with the recently proposed FSM (Shnain *et al.*, 2017). Figure 5 shows confidence in face recognition under noise using similarity threshold of *k* = 0.3, while using a higher threshold makes the confidence measure more relaxed.

Figure 6 shows the face recognition under blur (without noise) using similarity measure and AT&T database. Two values for the spread (standard deviation) of the blur Gaussian filter are used. Figure 7 shows the confidence in face recognition under blur using similarity threshold of *k* = 0.3 versus different values of blur filter spread (STD).

Only the proposed SVWG survives under the strict conditions of the 3rd confidence measure. Results confirm superiority of SVWG after examining the face recognition capabilities under blur - noise conditions using similarity measures.

Figure 8 depicts the face recognition under haze (without noise) using similarity measures for two values of haze level. Note that at higher levels of haze, all similarity measures might give higher similarity with non-similar images (or even fail, except for SSIM and SVWG), hence reducing recognition confidence. Figure 9 shows confidence in face recognition under haze (without no noise) using similarity threshold of *k* = 0.3. Under haze and strict confidence conditions, only the proposed SVWG survives.

Future Directions

Future attempts will include testing of the proposed SVWG under impulsive noise and using different Gabor filters. Tests will also handle transmission over real channel models (Mahmoud *et al.*, 2002; 2006) and incorporate Viola-Jones algorithm for better accuracy (Viola and Jones, 2001). Testing over OFDM, chaotic communication and Nakagami Channels is planned (Lau and Hussain, 2005; Al-Hinai *et al.*, 2007; Gurung *et al.*, 2010; 2008). Local analysis would also be considered due to its efficiency in handling facial expressions (Lajevardi and Hussain, 2009b). Transformational approaches in the Zernike and other domains could also be handled (Ali *et al.*, 2017).
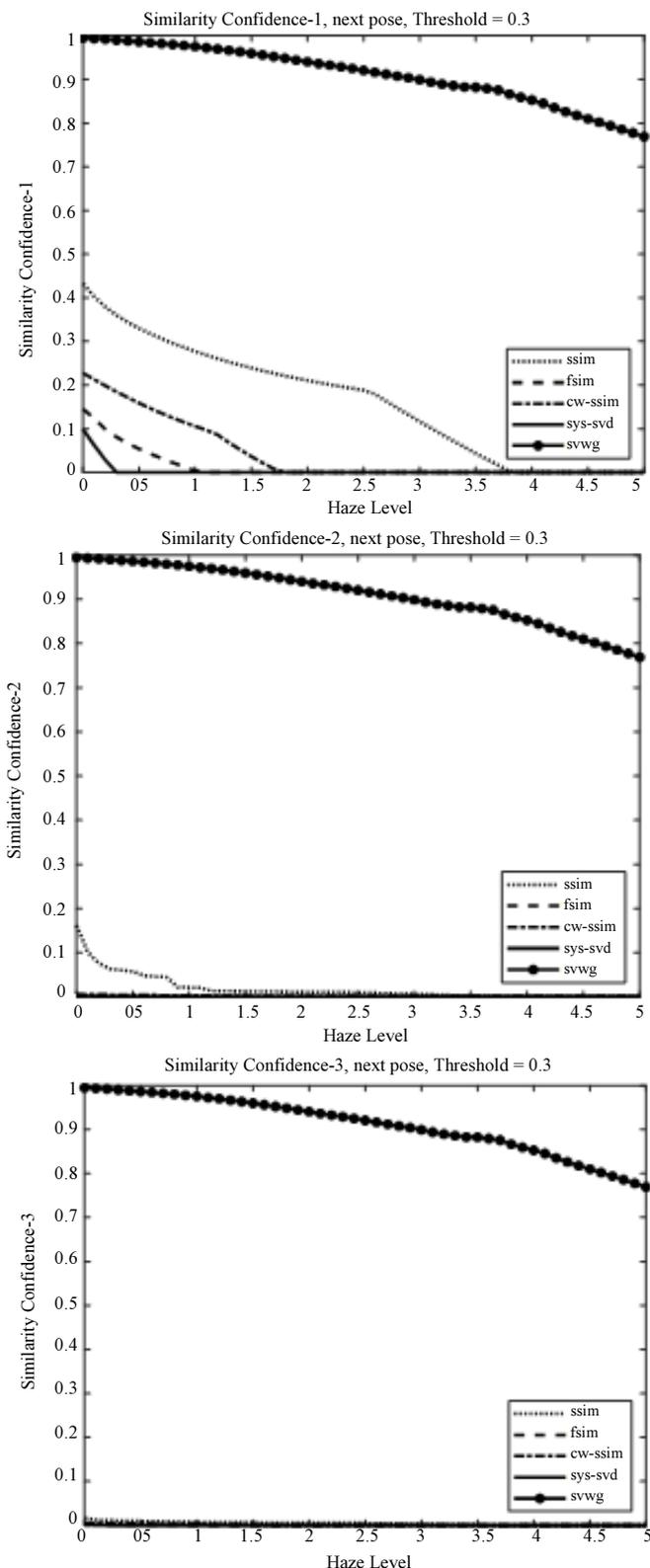
Fig. 9: Confidence in face recognition under haze (no noise) using similarity threshold of 0.3. Three confidence measures have been used. Note that a different pose is used for testing recognition confidence, as in Figure 8. Left, middle, and right figures represent the three confidence measures versus haze level as shown. Note that the 3rd measure is stronger than the 2nd, which is in turn stronger than the first. Only the proposed svwg survives under the 3rd confidence measure

Conclusion

This work presents an image similarity measure that is efficient in face recognition in the sense that it gives maximal similarity when the two face images under test are related to the same person, while it gives minimal amount of similarity when the two images are different. A shortcoming of existing similarity measures is that they may give a significant amount of similarity between unrelated images; the reason behind this fact is that each measure focuses on one direction in feature extraction, while the image is a complicated structure of properties. The proposed measure extracts features related to four different properties of the image: ICA structure, statistical properties (mean, variance and co-variance), frequency content and textural structure. In addition to ICA via SVD breakdown, the well-known similarity measure SSIM is used to reveal structural properties in the wavelet and Gabor domains, who in turn reveal spectral and textural properties of the image. Numerical results have shown significant improvement in face recognition, where higher confidence is shown when using the proposed measure (named as SVWG). Putting more conditions on recognition confidence, two new confidence measures have been defined. It is shown that, under the new confidence measures, the proposed SVWG outperforms existing similarity measures by far. Similar performance is found while more tests have been performed under adverse conditions of different pose, noise, blur and haze, where a new haze formation approach has also been proposed.

Acknowledgement

Author's Contributions

This is the MSc project of Ms. Ohood F. Ismael. She has contributed to the analysis and simulation of the proposed measures. Also, she contributed to the paper write-up. Zahir M. Hussain is the supervisor of this project. He has contributed to the analysis and simulation of the proposed measures. Also, he contributed to the write-up and language revision.

Ethics

The Authors declare that there are no ethical issues associated with this work.

References

Gurung, A.K., Al-Qahtani, F. S., Hussain, Z. M. and Alnuweiri, H., 2010. Performance analysis of amplify-forward relay in mixed Nakagami-m and Rician fading channels. Proceedings of the International Conference Advanced Technologies Communications, Oct. 20-22, IEEE Xplore press, Vietnam. DOI: 10.1109/ATC.2010.5672698

AT&T, 2019. The Database of Faces" (formerly "The ORL Database of Faces. AT&T Laboratories Cambridge, which can be accessed on http://www.face-rec.org/databases/.

Ali, A.J., Z.M. Hussain, M.S. Mechee and Z.A. Khalaf, 2017. Schoenberg logarithmic image similarity in prewitt-gabor-zernike domain. Am. J. Appli. Sci.

Gurung, A. K., Al-Qahtani, F. S., Sadik, A. Z. and Hussain, Z. M., 2008. Power savings analysis of clipping and filtering method in OFDM systems. Proceedings of the Australasian Telecommunication Networks and Applications Conference, Dec. 7-10, IEEE Xplore press, Australia. DOI: 10.1109/ATNAC.2008.4783323

Hashim, A.N. and Hussain, Z. M., 2014. Novel image-dependent quality assessment measures. J. Comput. Sci., 10: 1548-1560. DOI: 10.3844/jcssp.2014.1548.1560

Hassan, A.F., D. Cai-lin and Z.M. Hussain, 2014. An information-theoretic image quality measure: Comparison with statistical similarity. J. Comput. Sci., 10: 2269-2283. DOI: 10.3844/jcssp.2014.2269.2283

Chalom, E., E. Asa and E. Biton, 2013. Measuring image similarity: An overview of some useful applications. Proceedings of the IEEE Instrumentation Measurement Magazine, 16: 24-28. DOI: 10.1109/MIM.2013.6417053

Gabor, D., 1946. Theory of communication. J. Institution Electrical Engineers.

Park, D., H. Park, D.K. Han and H. Ko, 2014. Single image dehazing with image entropy and information fidelity. Proceedings of the IEEE International Conference Image Processing, (ICI'14).

Sang, G., J. Li and Q. Zhao, 2016. Pose-invariant face recognition via RGB-D images. Comput. Intelligence Neurosci., 2016: 1-9. DOI: 10.1155/2016/3563758 2016

Daugman, J. G., 1985. Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters. Optical Society Am.

He, K., Sun, J. and Tang, X. 2011. Single image haze removal using dark channel prior. I. Transact. Pattern Analysis Machine Intel., 33: 2341-2353. DOI: 10.1109/TPAMI.2010.168

Zhang, L., Zhang, Lei, X. Mou and D. Zhang, 2011. FSIM: A feature similarity index for image quality assessment. IEEE Trans. Image Proc.

Slavković, M. and D. Jevtić, 2012. A Face Recognition System Based on Eigenfaces Method. Serbian J. Electrical Eng., 9: 121-130. DOI: 10.2298/SJEE1201121S

Turk, M.A. and A P. Pentland, 1991. Face recognition using eigenface method. Proceedings of the IEEE Conference Computer Vision Pattern Recognition.

Sampat, M.P., Z. Wang, S. Gupta, A.C. Bovik and M.K. Markey, 2009. Complex wavelet structural similarity: A new image similarity index. IEEE Trans. Image Proc.

Aljanabi, M.A., Z.M. Hussain and S.F. Lu, 2018. An entropy-histogram approach for image similarity and face recognition. Math. Problems Eng. (Hindawi), 2018: 18-18. DOI: 10.1155/2018/9801308

Çarıkçı, M. and Özen, F. 2012. A face recognition system based on eigenfaces method. Proc. Technol., 1: 118-123. DOI: 10.1016/j.protcy.2012.02.023

Shnain, N.A., Z.M. Hussain and S.F. Lu, 2017. A feature-based structural measure: An image similarity measure for face recognition. Appli. Sci., 7: 786-786. DOI: 10.3390/app7080786

Al-Hinai, N., Neville, K., Sadik, A. Z. and Hussain, Z. M. 2007. Compressed image transmission over FFT-OFDM: A Comparative Study. Proceedings of the Australasian Telecommunication Networks Applications Conference, Dec. 2-5, IEEE Xplore press, Christchurch.

Viola, P. and M. Jones, 2001. Rapid object detection using a boosted cascade of simple features. Proceedings of the IEEE Computer Society Conference Computer Vision Pattern Recognition.

Jafri, R. and Arabnia, H. R. 2009. A survey of face recognition techniques. J. Information Proc. Syst., 5: 41-68.

Mahmoud, S. S., Hussain, Z. M. and O'Shea, P. 2002. Geometrical model for mobile radio channel with hyperbolically distributed scatterers. Proceedings of the 8th International Conference Communication Systems, Nov. 28, IEEE Xplore press, Singapore, DOI: 10.1109/ICCS.2002.1182428

Mahmoud, S.S., Z.M. Hussain and P. O'Shea, 2006. A geometrical-based microcell mobile radio channel model. Wireless Networks.

Lajevardi, S.M. and Z.M. Hussain, 2009a. Facial expression recognition: Gabor filters versus higher-order correlators. Proceedings of the International Conference Communication, Computer Power, Feb. 15-18, IEEE Xplore press, Muscat, Oman, pp: 1-5.

Lajevardi, S.M. and Z.M. Hussain, 2009b. Local feature extraction methods for facial expression recognition. P Proceedings of the 17th European Signal Processing Conference, Aug. 24-26, IEEE Xplore press, Glasgow.

Lajevardi, S.M. and Z.M. Hussain, 2009c. Zernike moments for facial expression recognition. Proceedings of the International Conference Communication, Computer Power, Feb. 15-18, Muscat, Oman,

Mahto, S. and Y. Yadav, 2014. A survey on various facial expression recognition techniques. IET Biometrics.

Singh, S. and S.V.A.V. Prasad, 2018. Techniques and challenges of face recognition: A critical review. Proc. Comput. Sci., 143: 536-543. DOI: 10.1016/j.procs.2018.10.427

Lee, S., S. Yun, J.H. Nam, C.S. Won and S.W. Jung, 2016. A review on dark channel prior based image dehazing algorithms. EURASIP J. Image Video Proc., Springer.

The MathWorks, 2019. MATLAB Documentation, accessible on https://nl.mathworks.com/help/ .

Zhang, Y.J., 2011. Advances in face image analysis. Publisher: Medi. Information Sci. Ref., NY, USA.

Lau, Y.S. and Z.M. Hussain, 2005. A new approach in chaos shift keying for secure communication. Proceedings of the IEEE International Conference Information Technology Applications, Jul. 4-7, IEEE Xplore press, Australia. DOI: 10.1109/ICITA.2005.30

Hussain, Z. M. and Boashash, B. 2001. Design of time-frequency distributions for amplitude and IF estimation of multicomponent signals. Proceedings of the Its International Symposium Signal Applications, Aug. 13-16, IEEE Xplore press, Kuala Lumpur, Malaysia.

Wang, Z., A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, 2004. Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Proce., 13: 600-612. DOI: 10.1109/TIP.2003.819861