[1.]S. HEMAJOTHI, [2.]S. ABIRAMI, [3.]S. AISHWARYA

A NOVEL COLOUR TEXTURE BASED FACE SPOOFING DETECTION USING MACHINE LEARNING

[1-3.]Department of ECE, Prince Shri Venkateshwara Padmavathy Engineering College, Chennai, INDIA

Abstract: Research on noninvasive software's based on face spoofing detection schemes has mainly been given focus on the analysis of the luminance information of the face images, hence removing the chroma component can be very much useful for discriminating photo faces from genuine faces. This brings us a novel and an appealing approach for detecting face spoofing using color texture analysis. The combined texture information and data of colors from the luminance and the chrominance channels by taking out complementary low-level feature descriptions from different color spaces are used. The color local binary patterns (LBP) descriptors, finds the facial color texture content using four other descriptors: the co-occurrence of adjacent local binary patterns (CoALBP), local phase quantization (LPQ). Here by using these features and characteristics the color texture is analyzed and extracted by face descriptors from different color bands. To gain insight into which color spaces are most suitable for discriminating genuine faces from fake ones, considered three color spaces, namely RGB, HSV and YCbCr. A new and appealing approach using color texture analysis and demonstrate that the chroma component can be very useful in discriminating fake faces from genuine ones. First, the face is detected, cropped and normalized into an M×N pixel image. Then, holistic texture descriptions are extracted from each color channel and the resulting feature vectors are concatenated into an enhanced feature vector in order to get an overall representation of the facial color texture. The final feature vectors are given to a binary classifier and the output score value describes whether it is a real or a fake image.
Keywords: spoofing detection schemes, noninvasive software, luminance information

INTRODUCTION

Compared with traditional authentication approaches including password, verification code and secret question, biometrics authentication is more user-friendly. Since the human face preserves rich information for recognizing individuals, face becomes the most popular biometric cue with the excellent performance of identity recognition. Currently, person identification can easily use the face images captured from a distance without physical contact with the camera on the mobile devices, e.g. mobile phone. As the application of face recognition system becomes more and more popular with the widespread of the Mobile phone, their weaknesses of security become increasingly conspicuous. For example, owing to the popularity of social network, it is quite easy to access a person's face image on the Internet to attack a face recognition system. Hence, a deep attention for face spoofing detection has been drawn and it has motivated great quantity of studies in the past few years. In general, there are mainly four types of face spoofing attacks: photo attack, masking attack, video replay attack and 3D attack. Due to the high cost of the masking attack and 3D attack, therefore, the photo attack and video reply attack are the two most common attacks. Photo and video replay attacks can be launched with still face images and videos of the user in front of the camera, which are actually recaptured from the real ones. Obviously, the recaptured image is of lower quality compared with the real one in the same capture conditions. The lower quality of attacks can result from: lack of high frequency information image banding or more effects, video noise signatures, etc. Clearly, these image quality degradation factors can work as the useful cues to distinguish the real faces and the fake ones. Face spoofing detection, which is also called face liveness detection, has been designed to counter different types of spoofing attacks.

The real and fake faces can be very distinctive in the chrominance channels. For instance, Burkinabe analyzed the impact of different color spaces on face anti-spoofing and presented a shallow model based on color features, achieving fairly good performance. To exploit and combine the effectiveness of color and deep learning, we utilize a deep learning framework combined with LBP features extracted from different color spaces (such as RGB, HSV, YCbCr). Among the significant contributions of our present work:

- While most previous methods based on deep learning suffer from the lack of face training samples, we introduce a new deep learning model by fine tuning the VGG-face model.
- We combine deep learning with handcrafted features by extracting the LBP descriptions from the convolutional feature maps.
- We explore how well different color spaces can be used for describing the intrinsic disparities in deep learning between genuine faces and fake ones. We also perform a fusion study to analyze the complementarity of different color space.

BINARY PATTERNS AND FACE QUANTAIZATION ALGORITHM

Here Texture descriptors originally designed for gray-scale images can be applied on color images by combining the features extracted from different color channels. Color texture of the face images is analyzed using three descriptors: Local Binary Patterns (LBP), Co-occurrence of Adjacent Local Binary Patterns (CoALBP), Local Phase Quantization (LPQ) have shown to be effective in gray-scale texture-based face anti-spoofing.

— LBP (Local Binary Pattern)

Local Binary Pattern (LBP) is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Parameters: the LBPH uses 4

The first computational step of the LBPH is to create an intermediate image that describes the original image in a better.

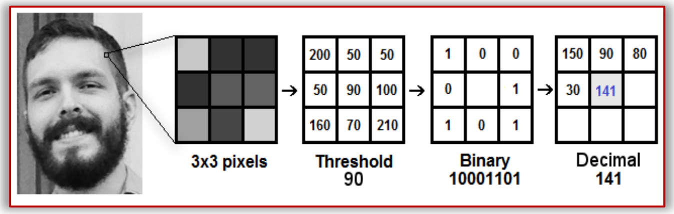


Figure 1. LBP image

For each neighbor of the central value (threshold), we set a new binary value. We set 1 for values equal or higher than the threshold and 0 for values lower than the threshold. Now, the matrix will contain only binary values (ignoring the central value) figure 1. We need to concatenate each binary value from each position from the matrix line by line into a new binary value (e.g. 10001101). Note: some authors use other approaches to concatenate the binary values (e.g. clockwise direction), but the final result will be the same. Then, we convert this binary value to a decimal value and set it to the central value of the matrix, which is actually a pixel from the original image. At the end of this procedure (LBP procedure), we have a new image which represents better the characteristics of the original image. The LBP procedure was expanded to use a different number of radius and neighbors, it is called Circular LBP. It can be done by using bilinear interpolation. If some data point is between the pixels, it uses the values from the 4 nearest pixels (2x2) to estimate the value of the new data point.

— LPQ (Local Phase Quantization)

In recent years, high elongation materials are widely used. Therefore, it is important to investigate the tensile properties of high elongation materials for engineering applications. Video extensometer is equipment for measuring the materials' tensile properties. It uses image processing technology to match data points and measures the actual deformation. However, when measuring high elongation materials, motion blur will appear on the collected images, which can affect the accuracy of image matching. In this paper, we proposed an image matching method which is based on Local Phase Quantization (LPQ) features to reduce the interference of the motion blur and improve the accuracy of the image matching algorithms as well. The experimental results on simulations show that the proposed initialization method is sufficiently accurate to enable the correct convergence of the subsequent optimization in the presence of motion blur. The experiment of uniaxial tensile also verifies the accuracy and robustness of the method. High elongation materials are an important class of materials for structural applications such as transportation, civil infrastructures, and biomedical applications. In actual service conditions, these materials are often subject to both mechanical and environmental loads. These factors will change the material properties and thus have a great influence on the service life and safety performance of these materials. In order to study these factors, the tensile mechanical test should be carried out on these materials. At present, the most commonly used equipment for the tensile mechanical test is the mechanical extensometer and video extensometer. For the high elongation materials, the mechanical extensometer which is mounted directly onto the material via blade causes many problems such as the following:

- mutual friction will reduce the measurement accuracy;
- the total deformation cannot be easily measured in the uniaxial tensile test;
- the measuring range is limited.

Compared with mechanical extensometer, the video extensometer has the following advantages. However, if higher accuracy is pursued, some influence factors cannot be ignored, such as out-of-plane displacement, self-heating of the camera, lens distortion, and image blur induced by motion. Reference [2] theoretically describes the measurement errors caused by out-of-plane displacement and self-heating of the camera; it further establishes a high-accuracy two-dimensional digital image correlation (2D-DIC) system using a bilateral telecentric lens to minimize the errors.

Reference [3] investigates the systematic errors due to lens distortion using the radial lens distortion model and in-plane translation tests; it finds out that the displacement and strain errors at an interrogated image point not only are linear proportion to the distortion coefficient of the camera lens used but also depend on the distance relative to distortion center and its magnitude of the displacement; the paper also proposes a linear least-squares algorithm to estimate the distortion coefficients and then to eliminate the errors.

It proposes an off-axis digital image correlation method for real-time, noncontact, and target less measurement of vertical deflection of bridges to achieve subpixel accuracy. Despite these advances, few works about eliminating extensometer's measurement errors caused by motion-induced image blur to improve the accuracy have been reported.

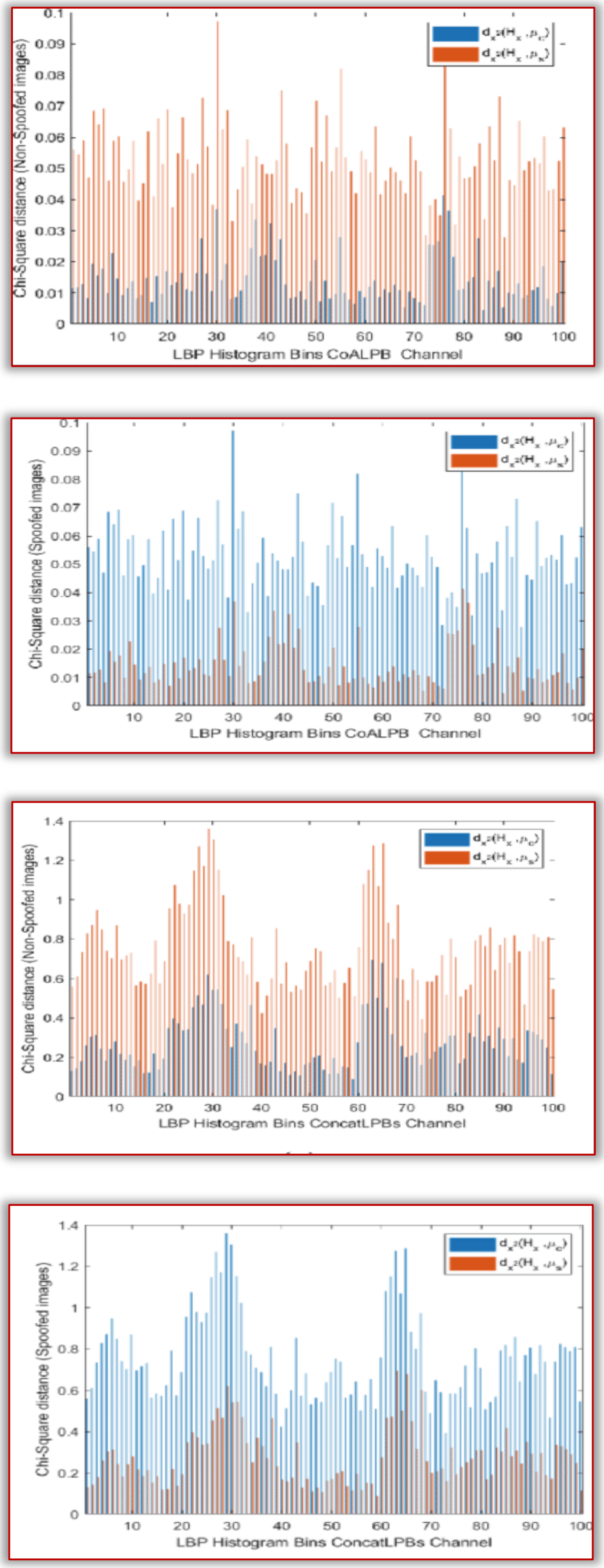
In this paper, we will propose an image matching method for video extensometer to measure the parameters by utilizing Local Phase Quantization (LPQ) feature. This method is robust and performs well on images with serious motion blur and deformation.

The phase information in the Fourier coefficients is recorded by examining the signs of the real and imaginary parts of each component in.This is done by using a simple scalar quantization, where is the component of the vector. The resulting eight binary coefficients are represented as integer values within 0–255 using binary coding:

Finally, a histogram is formed by all the positions in the rectangular region and used as a 256-dimensional feature vector in the match.

— CoLBP (Co-Occurrence Local Binary Pattern)

The new image feature based on spatial co-occurrence among micropatterns, where each micropattern is represented by a Local Binary Pattern (LBP). In conventional LBP-based features such as LBP histograms, all the LBPs of micropatterns in the image are packed into a single histogram in figure 2. Doing so discards important information concerning spatial relations among the LBPs, even though they may contain information about the image's global structure. To consider such spatial relations, we measure their co-occurrence among multiple LBPs. The proposed feature is robust against variations in illumination, a feature inherited from the original LBP, and simultaneously retains more detail of image. The significant advantage of the proposed method versus conventional LBP-based features is demonstrated through experimental results of face and texture recognition using public databases. Spatial co-occurrence could boost the discriminative power of the features, but it always suffers from the geometric and photometric variations. In this work, we investigate rotation invariant property of co-occurrence feature, and introduce a novel pairwise rotation invariant co-occurrence local binary pattern (PRI-CoLBP) feature which incorporates two types of context, spatial co-occurrence and orientation co-occurrence. Different from traditional rotation invariant local features, pairwise rotation invariant co-occurrence features preserve the relative angles between the orientations of individual features. The relative angle depicts the local curvature information, which is discriminative.



(a)

(b)

(c)

(d)

Figure 2. CoLBP image

METHODOLOGY

— Input image

Read and Display an input Image. Read an image into the workspace, using the misread command. In image processing, it is defined as the action of retrieving an image from some source, usually a hardware-based source for processing. It is the first step in the workflow sequence because, without an image, no processing is possible. The image that is acquired is completely unprocessed.

— **RGB colour image**

The RGB color model is an additive color model in which red, green, and blue light are added together in various ways to reproduce a broad array of colors. The name of the model comes from the initials of the three-additive primary colors, red, green, and blue.

The main purpose of the RGB color model is for the sensing, representation, and display of images in electronic systems, such as televisions and computers, though it has also been used in conventional photography. Before the electronic age, the RGB color model already had a solid theory behind it, based in human perception of colors' is a device-dependent color model: different devices detect or reproduce a given RGB value differently, since the color elements (such as phosphors or dyes) and their response to the individual R, G, and B levels vary from manufacturer to manufacturer, or even in the same device over time in figure 3. Thus an RGB value does not define the same color across devices without some kind of color management. Typical RGB input devices are color TV and video cameras, image scanners, and digital cameras. Typical RGB output devices are TV sets of various technologies (CRT, LCD, plasma, etc.), computer and mobile phone displays, video projectors, multicolor LED displays, and large screens such as Jumbotron. Color printers, on the other hand, are not RGB devices, but subtractive color devices (typically CMYK color model).

Figure 3. RGB image

— **Grayscale**

In photography and computing, a grayscale or greyscale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest. Grayscale images are distinct from one-bit bi-tonal black-and-white images, which in the context of computer imaging are images with only the two colors, black, and white (also called bilevel or binary images). Grayscale images have many shades of gray in between. Grayscale images are also called monochromatic, denoting the presence of only one (mono) color (chrome). Grayscale images are often the result of measuring the intensity of light at each pixel in a single band of the electromagnetic spectrum (e.g. infrared, visible light, ultraviolet, etc.), and in such cases they are monochromatic proper when only a given frequency is captured. But also, they can be synthesized from a full colour image; see the section about converting to grayscale in figure 4. Example of gray scale image is given below.

Figure 4. Gray scale image

— **Colour convertion**

A color space is a specific organization of colors. In combination with physical device profiling, it allows for reproducible representations of color, in both analog and digital representations. A color space may be arbitrary, with particular colors assigned to a set of physical color swatches and corresponding assigned color names or numbers such as with the Pantone collection, or structured mathematically as with the NCS System, Adobe RGB and sRGB. A "color model" is an abstract mathematical model describing the way colors can be represented as tuples of numbers (e.g. triples in RGB or quadruples in CMYK); however, a color model with no associated mapping function to an absolute color space is a more or less arbitrary color system with no connection to any globally understood system of color interpretation. Adding a specific mapping function between a color model and a reference color space establishes within the reference color space a definite "footprint", known as a gamut, and for a given color model this defines a color space. For example, Adobe RGB and sRGB are two different absolute color spaces, both based on the RGB color model. When defining a color space, the usual reference standard is the CIELAB or CIEXYZ color spaces, which were specifically designed to encompass all colors the average human can see. Since "color space" identifies a particular combination of the color model and the mapping function, the word is often used informally to identify a color model. However, even though identifying a color space automatically identifies the associated color model, such a usage is incorrect in a strict sense. For example, although several specific color spaces are based on the RGB color model, there is no such thing as the singular RGB color space.

— **Pre-processing**

Pre-processing is a common name for operations with images at the lowest level of abstraction both input and output are intensity images.

The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing. Image pre-processing methods use the considerable redundancy in images. Neighboring pixels corresponding to one object in real images have essentially the same or similar brightness value. Thus, distorted pixel can often be restored as an average value of neighboring pixels.

— **Feature extraction**

In machine learning, pattern recognition and in image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is related to dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images presented as pixels), then it can be transformed into a reduced set of features (also named a feature vector). Determining a subset of the initial features is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.

- ≡ Shape features
- ≡ color features

— **Classification**

Image classification refers to the task of extracting information classes from a multiband raster image. The resulting raster from image classification can be used to create thematic maps. The recommended way to perform classification and multivariate analysis is through the Image Classification toolbar. There are many classification algorithms are available and some classification algorithm that are given below.

In pattern recognition, the k-nearest neighbors' algorithm (k-NN) is a non-parametric method used for classification and regression.[1] In both cases, the input consists of the k closest training examples in the feature space. The output depends on whether k-NN is used for classification or regression.

In k-NN classification, the output is a class membership. An object is classified by a plurality vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of that single nearest neighboring k-NN regression, the output is the property value for the objecting figure 5. This value is the average of the values of k nearest neighbors. k-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. Both for classification and regression, a useful technique can be to assign weights to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones. For example, a common weighting scheme consists in giving each neighbor a weight of 1/d, where d is the distance to the neighbor.

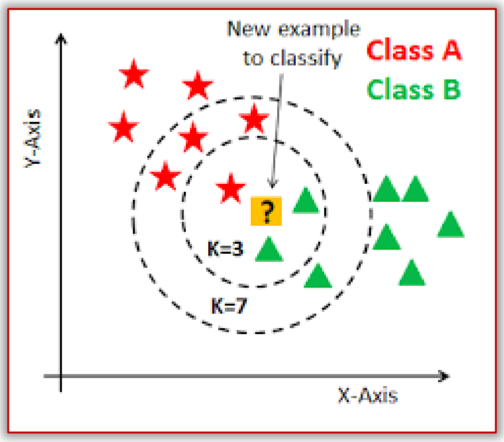


Figure 5. KNN

CONCLUSIONS

To approach the problem of face anti-spoofing from the color texture analysis point of view. We investigated how well different color image representations (RGB, HSV and YCbCr) can be used for describing the intrinsic disparities in the color texture between genuine faces and fake ones and if they provide complementary representations. The effectiveness of the different facial color texture representations was studied by extracting different local descriptors from the individual image channels in the different color spaces. The facial color texture representation seems be more stable in unknown conditions than texture descriptions extracted from gray-scale images. Thus, the use of color texture information provides a way to improve the unsatisfying generalization capabilities of texture-based approaches. In order to benefit from the potential complementarity of the CoALBP and the LPQ face descriptions, to fuse them by concatenating their resulting histograms. The facial representations extracted from different color spaces using different texture descriptors can also be concatenated in order to benefit from their complementarity. The effectiveness of different texture descriptors more closely in detecting various kinds of face spoofs by extracting holistic face representations from luminance and chrominance images in different color spaces. In future, the T test is use for feature selection to reduce high time consumption due to high feature size, second the Classifier need high training sample so by applying sparse classifier it will be reduced.

References

[1] Y. Li, K. Xu, Q. Yan, Y. Li, R.H. Deng, "Understanding OSN-based facial disclosure against face authentication systems," in Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security, ser. ASIA CCS '14. ACM, 2014
[2] J. Li, Y. Wang, T. Tan, A.K. Jain, "Live face detection based on the analysis of Fourier spectra," in Biometric Technology for Human Identification, 2004.
[3] X. Tan, Y. Li, J. Liu, L. Jiang, "Face liveness detection from a single image with sparse low rank bilinear discriminative model," in Proceedings of the 11th European conference on Computer vision: Part VI, ser. ECCV'10, 2010.
[4] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, S.Z. Li, "A face antispoofing database with diverse attacks," in 5th IAPR International Conference on Biometrics (ICB), 2012, pp. 26–31.
[5] J. Galbally, S. Marcel, "Face anti-spoofing based on general image quality assessment," in Proc. IAPR/IEEE Int. Conf. on Pattern Recognition, ICPR, 2014.
[6] D. Wen, H. Han, A. Jain, "Face spoof detection with image distortion analysis," Transactions on Information Forensics and Security, vol. 10.
[7] J. Bai, T.-T. Ng, X. Gao, Y.-Q. Shi, "Is physics-based liveness detection truly possible with a single image?" in IEEE International Symposium on Circuits and Systems (ISCAS), 2010.
[8] J. Mahatma,¨ A. Hadid, M. Pietikainen, "Face spoofing detection from single images using micro-texture analysis," in Proceedings of International Joint Conference on Biometrics
[9] J. Komulainen, A. Hadid, and M. Pietikainen, "Face spoofing detection from single images using texture and local shape analysis," Biometrics, IET, vol. 1, no. 1, pp. 3–10, March
[10] J. Yang, Z. Lei, S. Liao, S. Z. Li, "Face liveness detection with component dependent descriptor," in IAPR International Conference on Biometrics, ICB, June 2013.
[11] T. d. F. Pereira, J. Komulainen, A. Anjos, J. M. De Martino, A. Hadid, M. Pietikainen, S. Marcel, "Face liveness detection using dynamic texture", EURASIP Journal on Image and Video Processing, 2013

ACTA TECHNICA CORVINIENSIS – Bulletin of Engineering
ISSN: 2067-3809