CUBIC-SPLINES NEURAL NETWORK- BASED SYSTEM FOR IMAGE RETRIEVAL

Samy Sadek, Ayoub Al-Hamadi, Bernd Michaelis

IESK
Otto-von-Guericke University Magdeburg
Magdeburg, Germany

Usama Sayed

Faculty of Engineering
Assiut University
Assiut, Egypt

ABSTRACT

Research in Content-Based Image Retrieval (CBIR) shows that high-level semantic concepts in image cannot be constantly depicted using low-level image features. So the process of designing a CBIR system should take into account diminishing the existing gap between low-level visual image features and the high-level semantic concepts. In this paper, we propose a new architecture for a CBIR system named SNNIR (Splines Neural Network-based Image Retrieval). SNNIR system makes use of a rapid and precise neural model. This model employs a cubic-splines activation function. By using the spline neural model, the gap between the low-level visual features and the high-level concepts is minimized. Experimental results show that the proposed system achieves high accuracy and effectiveness in terms of precision and recall compared with other CBIR systems.

Index Terms— Cubic-splines neural network, feature extraction, content-based retrieval.

1. INTRODUCTION

A CBIR system mainly aims at avoiding the use of textual descriptions and instead retrieves images based on their visual similarity to a user-supplied query image or user-specified image features. Recent years have witnessed a fast growth of the use of digital image collections in many domains such as multimedia libraries, document archives, medical image management and biometrics. This growth motivates the research in image retrieval and brings the need for efficient CBIR techniques[1]. Loads of papers have been published in the last few years in this area. For instance, Wan and Kuo [2] have proposed an approach for image retrieval with hierarchical color clustering. In [3], a system that mimics the human brain for CBIR was presented. CBIR systems still do not carry out as well as their text counterparts [4]. Image retrieval systems have traditionally relied on annotations or captions associated with the images for indexing the retrieval system. The labor-intensive task of indexing and cataloging the images in these collections has conventionally been performed manually, a process that can be subjective and prone to errors.

The last few decades have caught sight of abundant advancements in the area of CBIR [1]. Although CBIR approaches have demonstrated success in moderately constrained domains including pathology, dermatology, chest radiology, and mammography, they have verified poor performance when applied to databases with a wide spectrum of imaging modalities, anatomies and pathologies [4], [5], [6], [7]. Image retrieval performance has shown comprehensible improvement by fusing the results of textual and visual techniques [8], [9].

In this paper, a new neural system for image retrieval called Spline Neural Network Image Retrieval (SNNIR) is presented. The proposed system utilizes an adaptive neural network model called spline neural network (SNN). The spline neural network enables the system to determine nonlinear relationship between the low-level visual image features and high-level semantic concepts in images. Results show that the proposed system is more effective and efficient to retrieve visual-similar images for a set of images with same conception.

The rest of the paper is organized as follows. In section 2, the architecture of the cubic-splines network is presented. Extracted features are described in section 3. Section 4 investigates the architecture of the SNNIR system. In section 5, the experimental results of the proposed system are reported. Concluding remarks are offered in section 6.

2. CUBIC-SPLINES NEURAL NETWORKS

The activation function simulates the correlation between the action potential of the inputs and the output of the neuron. Artificial Neural Networks (ANN) implementations are frequently using well-known activation functions like the sigmoidal functions. In this paper we employ an adaptive activation function for the hidden neurons out of a pool of standard functions called cubic-splines function to increase flexibility. In general, choosing a function as an activation function

This work is supported by Bernstein-Group (BMBF, FKZ: 01GQ0702), Forschungspraemie (BMBF-Förderung, FKZ: 03FPB00213) and Transregional Collaborative Research Centre SFB/TRR 62 "Companion-Technology for Cognitive Technical Systems" funded by the German Research Foundation (DFG).

should take into account these aspects: Simple implementation, fast computation, and the used function should be partially refinable. These requirements lead to the mathematical field of interpolating polynomials, a part of numerical analysis. Cubic-splines function consists of third degree polynomials. Each two consecutive data points (i.e. knots) are connected by a specific polynomial. For a data set $\{x_i\}$ of $n+1$ points, a cubic splines function is defined as

$$S(x) = \mathbf{C}_k s_k(x) = \mathbf{C}_k \sum_{i=0}^{3} s_{k,i}(x - x_k)^i \tag{1}$$

$$\forall x \in [x_k, x_{k+1}], k = 0, 1, \ldots, n-1.$$

where **C** is the concatenation operator and $s_{k,i}$ are the coefficients of the cubic-splines function.

To determine the above coefficients , a system with four equations has to be set up. This system can be constructed using the properties of the splines function (e.g. the interpolating property, twice continuous differentiable, etc). By combining Eq. (1) and the splines' properties we finally get the following matrix of equations (see [10] for more details).

$$\begin{pmatrix} s_{k,0} \\ s_{k,1} \\ s_{k,2} \\ s_{k,3} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\alpha & 0 & \alpha & 0 \\ 2\alpha & \alpha-3 & 3-2\alpha & -\alpha \\ -\alpha & 2-\alpha & \alpha-2 & \alpha \end{pmatrix} X \tag{2}$$

$$X = \begin{pmatrix} x_{k-1} \\ x_k \\ x_{k+1} \\ x_{k+2} \end{pmatrix}, k = 1, 2, \ldots, n-1.$$

In Eq.(2), α is the tension parameter of cubic-splines function that determines how sharply the curve bends at the control points. The common value for α is 0.5. Therefore, if we set $\alpha = 0.5$ into Eq.(2), the cubic-splines function given in Eq.(1) can be determined as follows:

$$s_k(u) = \frac{1}{2} \begin{pmatrix} 1 & u & u^2 & u^3 \end{pmatrix} AX \tag{3}$$

$$A = \begin{pmatrix} 0 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 2 & -5 & 4 & -1 \\ -1 & 3 & -3 & 1 \end{pmatrix},$$

$$0 \le u \le 1.$$

3. EXTRACTED FEATURES

RGB color space is still not a satisfactory choice for image retrieval systems. Therefore all computations of the features are done using HSV color space. Feature vectors, which characterize images, are made up of the following color descriptors:

3.1. Color coherence

Although color histograms are frequently used in many applications, they have some limitations. For example they provide no spatial information; they only describe which colors are present in the image, and in what quantities. Moreover, color histograms are sensitive to compression artifacts and changes in overall image brightness. On the other hand, Color Coherence (CC) ,which is defined as the degree to which pixels of that color are members of large similarly-colored regions, is a different way of incorporating spatial information into the color histogram. The significant regions are termed as coherent regions, and they are of significant importance in characterizing images. In this way, each histogram bin is divided into two types: coherent, if it belongs to a large uniformly-colored region, or incoherent, if it does not. Let α_i, β_i denote the number of coherent and incoherent pixels in the *i-th* bin respectively. Then, the CC vector is defined as:

$$\langle (\alpha_1, \beta_1), (\alpha_2, \beta_2), \ldots, (\alpha_n, \beta_n) \rangle \tag{4}$$

3.2. Color moments

Color moments originate from the hypothesis that color distribution in an image can be interpreted as a probability distribution. Probability distributions are characterized by a number of unique moments (e.g. Normal distributions are discriminated by their mean and variance). Hence, if the color in an image follows a certain probability distribution, the moments of that distribution can be used as features to identify that image. The three central moments of an image are mean, standard deviation and skewness. Mathematically, these moments can be defined as:

$$\mu_k = \frac{1}{n} \sum_{i=1}^{n} p_i^k \tag{5}$$

$$\sigma_k = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (p_i^k - \mu_k)^2} \tag{6}$$

$$s_k = \sqrt[3]{\frac{1}{n} \sum_{i=1}^{n} (p_i^k - \mu_k)^3} \tag{7}$$

where p_i^k is the value of the *k-th* color channel at the *i-th* pixel, and n is the size of image.

4. SPLINES-NN IMAGE RETRIEVAL SYSTEM

In this section, the structure of the proposed SNNIR system is described. Fig. 1 shows the main components of the proposed system. There are two stages of the control flow. The first deals with learning the splines NN and saving feature vectors, while the second deals with the query image to retrieve all the similar images. During the learning stage, a set of images in the database has been grouped into predefined categories. The main components of the SNNIR system depicted in Fig.1 are briefly described as shown in following subsections.

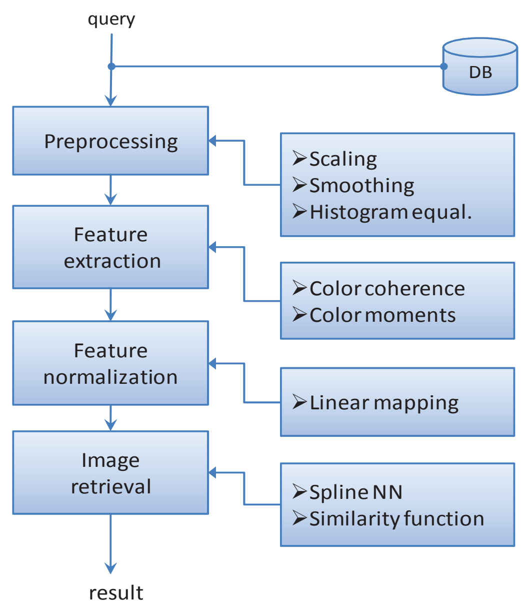

Fig. 1. Block diagram of the SNNIR system.

4.1. Pre-processing

A preliminary step of creating a symbolic representation of the source images is required before applying any data retrieval methods. The images are thus normalized by bringing them to a common resolution, performing histogram equalization and applying Gaussian filter to remove small distortions without reducing the sharpness of the image.

4.2. Feature extraction

Feature extraction is the basis of a CBIR system. Color is one of the most recognizable features of image. The main difficulty with any feature extraction operation is that the pixel is the unit of information in an image and each pixel has only two properties of position and color value. These knowledge of the position and value of a particular pixel should generally convey all information related to the image contents [11]. To avoid this difficulty, image features are extracted using two ways (i.e. color coherence and color moments). This allows the system to extract from an image a set of numerical values, expressed as coded characteristics of the image objects, and used to differentiate one category from another.

4.3. Feature normalization

This step is used to prevent singular features from dominating the others and to obtain comparable value ranges. Here we normalize the features by a linear scaling procedure that provides for all the features to have values in a unit range. Let the lower and upper bounds for a feature component, x are a, b respectively. Then, the linear scaling is given as:

$$\varphi(x) = \frac{x-a}{b-a} \tag{8}$$

4.4. SNN image retrieval

In this step, the two feature vectors of the query image and an image in the database are fed into the SNN. Then the system predicts the corresponding metric vectors which will be used to determine the similarity function. In this paper, we have used a cosine similarity metric function.The cosine similarity function between two images; an image in the database, I and the query image, q is defined as follows:

$$S(I,q) = \cos(\vec{v_I}, \vec{v_q}) = \frac{\vec{v_I} \cdot \vec{v_q}}{\|\vec{v_I}\|\|\vec{v_q}\|} \tag{9}$$

where $\vec{v_I}$ and $\vec{v_q}$ are the two metric vectors for the retrieval image and the query image respectively. Furthermore, in this step, user can provide the system a feedback which helps to update the spline activation function.

5. EXPERIMENTAL RESULTS

In this section, the experimental results of the proposed SNNIR system are presented. The designed cubic-splines network has three layers: input layer, hidden layer and output layer including 90, 20, 5 units respectively. The feature vector has 90 dimension containing 72-dimensional color coherence and 18-dimensional color moments. To evaluate the learning capabilities of the designed neural model, a comparison with standard sigmoidal neural modal has been carried out. The runs consist of training the two different models in the same retrieval problem. Fig. 2 shows the plots of the Mean Squared Error (denoted as MSE) during the learning stage.

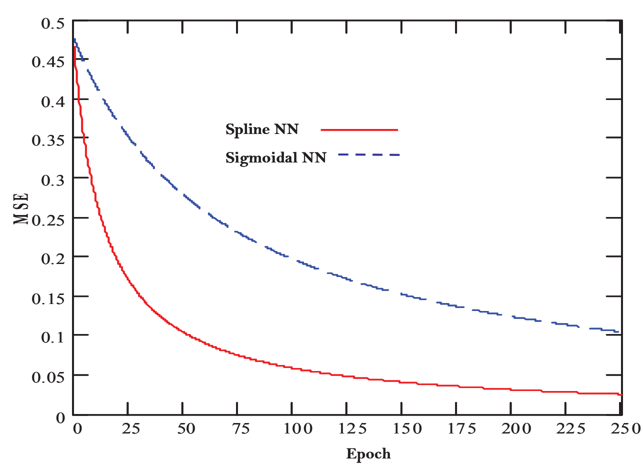

Fig. 2. Averaged learning curves comparison between cubic-splines NN and standard sigmoidal NN.

To verify the retrieval performance for the proposed SNNIR system, a general-purpose image database including 500 images has been used. The images in the database fall into five categories: Building, Coast, Mountain, Car, and Farm. Each category depicts a distinct semantic topic. To evaluate the efficiency of the proposed system, an image from each category has been selected as a query image. In addition, the effectiveness of the system has measured using two metrics: precision and recall. Table 1 shows different retrieval results for each query image obtained by the system. The figures in table 2 show that the SNNIR system performs well in the terms of precision and recall compared with other CBIR systems. For example, The proposed system outperforms the CBIR systems proposed in [12] and [13]. Furthermore, the proposed system is very efficient for the images of the same conception (see Fig. 3).

Table 1. The results of retrieval for query images.

Class	**Precision**	**Recall**
Building	0.89	0.96
Coast	1.00	1.00
Mountain	0.85	0.92
Car	0.90	0.72
Farm	1.00	0.92
Average performance	**0.93**	**0.90**

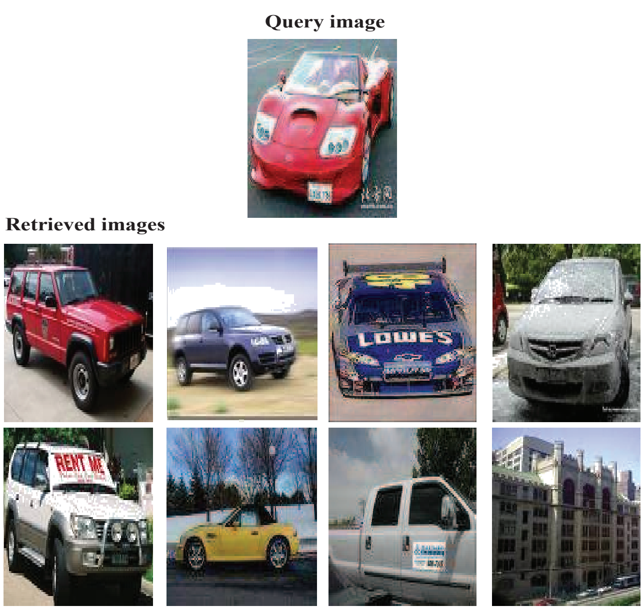

Fig. 3. Some retrieval results of SNNIR system.

6. CONCLUSION AND FUTURE WORK

In this paper, we have presented the SNNIR system. This system could successfully designate nonlinear relationships between low-level visual image features and the high-level semantic concepts in images. Our experimental results indicated that the SNNIR system discussed here is promising and bears further investigation and development. Our future work will include extending the system to deal with image databases of larger categories and looking into ways for improving the feature extraction process.

7. REFERENCES

[1] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. on Pattern Recognition and Machine Learning*, vol. 22, pp. 1349–1380, 2000.

[2] X. Wan and C.C.J. Kuo, "A new approach to image retrieval with hierarchical color clustering," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 8, pp. 628–643, 1998.

[3] A. Kulkarni, "CBIR using associative memories," in *Proceedings of the 6th WSEAS Int. Conference on Telecommunications and Informatics*. Dallas, Texas, USA, 2007, vol. II, pp. 99–104.

[4] W. Hersh, H. Muller, and et al, "Advancing biomedical image retrieval: development and analysis of a test collection," *J. Am. Med. Inform. Assoc.*, vol. 13, pp. 488–496, 2006.

[5] A. Aisen, L. Broderick, and et al., "Automated storage and retrieval of thin-section ct images to assist diagnosis:systemdescription and preliminary assessment," *Radiology*, vol. 228, pp. 265–270, 2003.

[6] P. Schmid-Saugeon, J. Guillod, and et al., "Towards a computer-aided diagnosis system for pigmented skin lesions," *Computerized Medical Imaging and Graphics*, vol. 27, pp. 65–78, 2003.

[7] H. Mller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval systems in medicine - clinical benefits and future directions," *Int. J. Med. Inform*, vol. 73, pp. 1–23, 2004.

[8] W. Hersh, J. Kalpathy-Cramer, and et al, "Effectiveness of global features for automatic medical image classification and retrieval," *Pattern Recognition Letters*, vol. 29, pp. 2032–2038, 2008.

[9] J. Kalpathy-Cramer and W. Hersh, "Automatic Image Modality Based Classification and Annotation to Improve Medical Image Retrieval," in *Proceedings of the 12th World Congress on Health (Medical) Informatics*, 2007, pp. 1334–1338.

[10] E. Catmull and R. Rom, "Computer aided geometric design, chapter a class of local interpolating splines," *academic Press*, pp. 317–326, 1974.

[11] R. Marmo et al, "Textural identification of carbonate rocks by image processing and neural network: Methodology proposal and examples," *Computers and Geosciences*, vol. 31, pp. 649–659, 2005.

[12] B. Si, W. Gao, H. Lu, and W. L. D. Zeng, "Image retrieval method based regions of interest," *High Technology Letters*, vol. 31, pp. 13–18, 2003.

[13] J.Z. Wang, J.Li, and G. Wiederhold, "SIMPLIcity: semanticsensitive integrated matching for picture libraries," *IEEE Trans. Image Processing*, vol. 9, pp. 947–963, 2000.