ISSN: 2321-7782 (Online)
Impact Factor: 6.047
Volume 4, Issue 6, June 2016
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

Sign Language Recognition Using Principal Component Analysis

P. Revathi
M.Sc.,
Department of Computer Science,
Auxilium College (Autonomous),
Vellore, Tamil Nadu – India

Abstract: Sign language recognition is an important research problem for enabling communication with blind impaired people. This paper presents principal component analysis which is a fast and efficient technique for recognition of sign gestures from video stream. Capturing of images from live video can be done using webcam or an android device. Proposed System is capable to capture three frames per second from video stream. After the Acquisition, system compares three continuous frames to know the frame, containing static posture shown by hand. This static posture is recognized as a sign gesture. Now it is matched with stored gesture database to know its meaning. It suggests an alternative path in voice or audio format and this system can be applied in real time.

Keywords: Recognition; Communication; Impaired; Gestures; Acquisition; Audio; Frames.

I. Introduction

Of the 314 million visually impaired peoples worldwide, 45 million are blind. Recognizing the objects in front of them is obviously essential today. Those people cannot navigate easily in their day to day life. They need help of others or cane or other electronic mobility devices or guided dogs which guide them in an appropriate manner. So they need a self assistive device to guide and make them independent from being dependent on others for navigation. The very preliminary and significant thing is the detection of obstacles in front of them and avoiding it. This paper needs to classify the objects, recognize obstacles and track the objects through an image processing technique called Principal Component Analysis (PCA) and suggests them an alternative path through voice format.

A. Objective

The main aim is to automatically extract foreground objects in videos which are captured by freely moving cameras. Instead of assuming that the background motion is dominant and different from that of the foreground, consider the assumption and allow foreground objects to be presented in freely moving scenes. Advancement of both visual and motion saliency information across video frames and a model is utilized for integrating the associated features for visual saliency, shape, foreground/background colour models, and spatial/temporal energy terms.

II. System Analysis

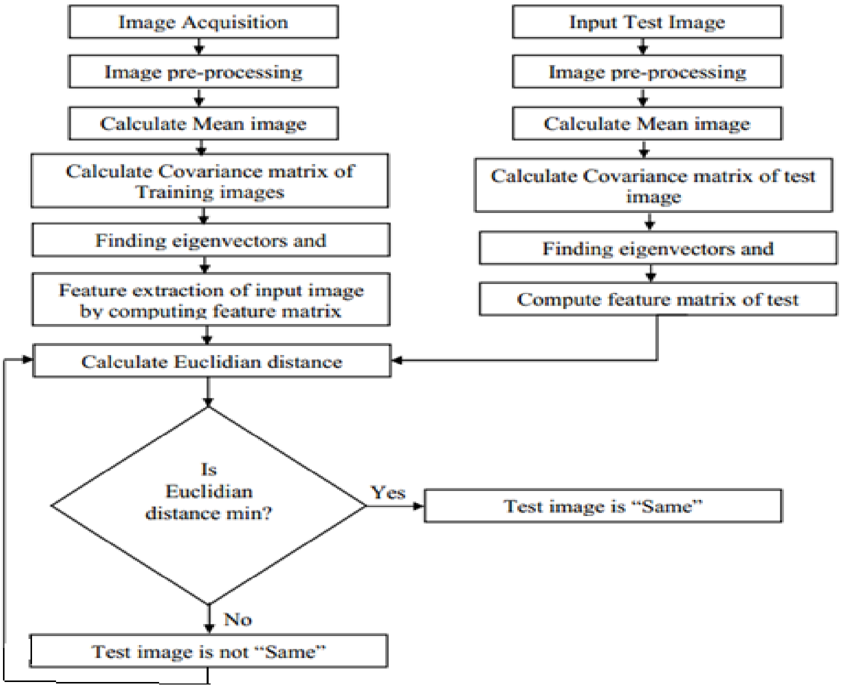
A. Existing System

Hand modelling and tracking are essential in video-based sign language recognition. The high reform ability and the large number of degrees of freedom of hands render the problem difficult. To tackle these challenges, a novel approach based on robust Principal Component Analysis (PCA) is proposed. The robust PCA incorporates an L_1 norm objective function to deal with background clutter and a projection pursuit strategy to deal with the lack of alignment due to the deformation of hands. The robust PCA, a hand tracking system is developed that contains a skin-colour region segmentation based on graph cuts and template matching in the framework of particle filtering.

B. Proposed System

Sign Language is the primary language used by the Blind community in order to convey information through gestures instead of words. In addition, this language is also used for human-computer interaction. This is an approach which can recognize sign language, based on principal component analysis and artificial neural network. Our approach begins by detecting the hand, pre-processing, determining Eigen space to extract features and using artificial neural network for training and testing. This method has low computational cost and can be applied in real-time. The proposed approach has been tested with high accuracy.

III. SYSTEM STRUCTURE



IV. SIGN LANGUAGES RECOGNITION

A. Image Resizing

This step will standardize the image size, prepare for training and recognition and increase the recognition accuracy. The standard resizing methods, the hand can be shrunk or stretched with a different (horizontal and vertical) ratio, so the hand characteristics and the recognition results are affected. The image must be cropped by the object's bounding box, where h and w are the height and width of the image, respectively. α is defined as the difference between w and h. Proposed method to adjust the hand size is as follows:

Resize the obtained image to 100*100 pixels using standard resizing methods.

B. Recognition

Feature Extraction: To extract the features from training images, Principal Component Analysis Algorithm is considered as below:

Step 1: Obtain a set S with M training images; Each image is transformed into a equal vector size as N =10000 (100*100) and placed into the set.

Step 2: Compute the mean image

Step 3: Find the difference between each training image and the mean image

Step 4: Obtain the co-variance matrix C in the following manner.

Step5: Find a set of M Ortho normal vector ui which describes the best distribution of the data. Since the size of the matrix C is too big (10000*10000), so to find ui, we seek eigenvectors vi of the matrix L = ATA with size M*M.

Step 6: Compute eigenvectors of the matrix C based on eigenvectors of the matrix L.

Step 7: Each training image Γ is transformed into new space and represented as a vector Ω.

C. Network Training and Recognition

Pattern recognition can be implemented by using a trained feed forward neural network. Multi layer Neural Network is a widespread disseminated effective method. In the proposed approach, the multilayer feed forward neural network is trained by the Back Propagation algorithm. Due its easy implementation, fast and efficient operation, multilayer feed forward network is widely used in numerous applications. When the network is trained, the input patterns are recognized and distinguished by the associated output patterns. The network progressively adjusts the outputs with respect to input patterns until approaching an error criterion. Subsequently, the best network architecture is selected with a test dataset. After testing, the final network delivers the output corresponding to the input pattern with minimal error.

D. Gesture Recognition

Gesture recognition is a topic in computer science and language technology with the goal of interpreting human gestures via mathematical algorithms. Gestures can originate from any bodily motion or state but commonly originate from the face or hand. Gesture recognition can be seen as a way for computers to begin to understand human body language, thus building a richer bridge between machines and humans. Gesture recognition enables humans to communicate with the machine and interact naturally without any mechanical devices. Gesture recognition can be conducted with techniques from computer vision and image processing.

Frame Extraction

Frame extraction is the first process of the system, this module reads the input video and extracts the number of frames from the corresponding video through an Android Mobile (AM). To compute texture amplitude, the intensity image is smoothed with a median filter and the result is subtracted from the original image. The absolute values of these differences are run through a second median filter.

Background Color Removal

Background colour removal is second process of the system to know recognition of the original object, this system constructs the colour histogram of each frame and removes the colour that appears most frequently in the scene. This module also removes the pixels which are not needed to consider in subsequent detection processing. Background colour removal not only reduces object information but also speeds up the detection Process.

Feature Extraction

This module is to extract the feature from the image frame. Process of feature extraction is Edge Detection, Corner Detection, Colour Transformation and Colour Classification and it will determine obstacle and image recognition as a result. Eigen space approaches are used for view-based approach. They provide an efficient representation of a large set of high dimensional points using a small set of orthogonal basis vectors. These basis vectors span a subspace of the training set called the Eigen space and a linear combination of these images can be used to approximately reconstruct any of the training images

Voice Convert

The voice convert module converts obstacle name which is already stored in mobile storage and GPS location with traffic signal colour value from detection speech synthesis is the artificial production of human speech. A computer system used for this purpose is called a speech synthesizer. A text-to-speech (TTS) system converts normal language text into speech; other systems render symbolic linguistic representations like phonetic transcriptions into speech. In this project, the speech synthesis employed is Window's text-to-speech – Microsoft Sam. The use of the existing speech synthesis is due to the nature and objective of this project, as emphasis is weighed on gesture recognition rather than text-to-speech.

Operational Components

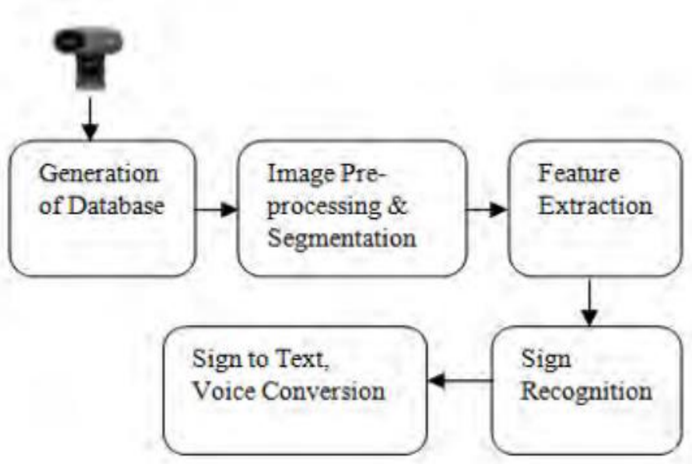
Android Smart phones any version from 2.3.3 gingerbread.

Emulator camera back side in device with Camera permission.

Edge Detection algorithm with segmentation based background elimination.

TTS activity of android mobile to enable voice translates about obstacle Status.

V. Principle of Sign Recognition



A. *Data Acquisition*

To achieve a high accuracy for sign recognition in sign language recognition system 260 images are used, 10 each of the 26 signs are used. These 260 images are included in training and testing database. The images are captured at a resolution of 3000x4000 pixels. The runtime images for test phase are captured using web camera. The images are captured in white background so as to avoid illumination effects. The images are captured at a specified distance (typically 1.5 – 2 feet) between camera and signer. The distance is adjusted by the signer to get the required image clarity.

B. *Image Pre Processing and Segmentation*

Preprocessing consists of image acquisition, segmentation and morphological filtering methods. Then the Segmentation of hands is carried out to separate object and the background. Otsu algorithm is used for segmentation purpose. The segmented hand image represents certain features. These features are further used for gesture recognition. Morphological filtering techniques are used to remove noises from images to get a smooth contour.

C. *Feature Extraction*

Feature extraction is a method of reducing data dimensionality by encoding related information in a compressed representation and removing less discriminative data. Feature extraction is vital to gesture recognition performance. Therefore, the selection of which features to deal with and the extraction method are probably the most significant design *decisions in hand motion and gesture recognition development. Centroid, skin colour and principal component are used as main features.*

D. Sign Recognition

It is a dimensionality reduction technique based on extracting the desired number of principal components of the multi-dimensional data. The gesture recognition using PCA algorithm involves two phases:

1) Training Phase

2) Recognition Phase

VI. Model-Based Tracking of Free Movable Objects

During the third step of the algorithm a virtual model of the movable object is inserted into the static 3D environment model. The position of this virtual object can be updated interactively according to the movements of the real object. If a virtual model corresponding 3D model of this object is available, it can be stored with the object description in the database. Otherwise, an abstract model or a simple bounding box can be used. When the real object is moved to a new location the corresponding location of the virtual object model can be updated interactively. This can be done by adding a new transformation node to the scene graph, which shifts the virtual object to the new location. This simultaneous tracking of the virtual object provides the blind user information about the real object, its features and its current location relative to the users. Further, it warns of obstacles in the current environment close to real-time demands.

VII. Training of Movable Object

In the first step of our algorithm, the object of interest is trained according to its shape and its colour. The corresponding object description and its name is stored in a database with shape images, colour histogram information and basic colour terms. For the segmentation of the outline, Alpha-blending between one of the colour images and the disparity image of the stereo camera is used. The colour image is multiplied by a factor of 0.9 and the disparity image is added with factor 0.1.

The resulting image is used for the segmentation of the region of interest using a standard region-growing algorithm. During the segmentation procedure the disparity information assures that only image parts with the correct depth are chosen. Otherwise objects in the background with the same colour could be selected too. For the selected region an image mask is generated and stored in a database together with the pertinent colour histogram and the basic colour term of this region, the colour histogram consists of 16 hue-ranges of the Hue Staturations Intensity (HIS) colour model. The basic colour term is determined with an algorithm that takes into account the region’s colour environment. This reflects human colour perception, which is influenced by surrounding colours. To allow increased independence of the viewing direction during the search, several image masks containing shape information can be stored for one object in the database.

VIII. Testing and Implementation

System testing involves unit testing, integration testing, white-box testing and black-box testing. Strategies for integration of software components into a functional product include the bottom-up strategy, the top-down strategy and sandwich strategy. Careful planning and scheduling are required to ensure that modules are available for integration into evolving it into a software product. When needed a series of testing are performed for the proposed system. Before the system is ready for the user, acceptance testing will be performed.

Table1 Testing and Implementation

Testing	Unit Test	Integration Test	Performance Test	Validation Test	Block Box Test	White Box Test
Voice Recognition	TTS test	Integration of TTS with application test	Output of TTS test	validate the voice of TTS	Select voice	TTS voice activity test
Live Video Capture	Camera test	Camera integration with application test	Output of camera test	Validate test the camera activity and capture flow	Camera intent test	Camera capture test
Object Detect	Testing train and test directory	Integration of matching object with application test	Output of matching case	Using recognition algorithm based test	Algorithm test	Detection test

IX. User Manual

A. Open the ADT Bundled Eclipse on start menu, import corresponding project folder and click run on the tool bar to execute the system.

B. After execution of the program, application will be generated by the ADT Bundled Eclipse as APK file that is the Application Programming Kit.

C. Then the APK file should be loaded into the Android Smart Phone and click the icon in the gadget as well as trigger on the service and background settings.

D. Click Track Object button on blind cam interact system then the automatic image acquisition could be performed.

E. After acquisition, captured image will be processed for background colour removal and feature extraction.

F. Extracted feature of original image is compared with trained images which are already stored in the mobile directory through Artificial Neural Network Technique.

If the testing obstacles are matched with the trained obstacles, means then its text is converted into audible voice format which is used by the visually impaired people.

X. Conclusion

In this system, contributed an algorithm and performed usability studies, which can extract high-quality and visually-rich frames from continuous camera video, experiments that evaluate and optimize the algorithm, an accessible real-time scanning application with which blind people can identify everyday objects around them. This approach works better than the current standard. Most camera interfaces lack accessibility for blind people even though many accessible mobile applications are picture-based. Scan Search improves blind users to experience accessibility in multiple areas.

Acknowledgement

I would like to thank the anonymous reviewers for their valuable and constructive comments on improving the paper. This work was supported by the Auxilium College, vellore of Tamil Nadu, the development of New Century Excellent Talents in individual Research persons.

References

1. Cooper and R. Bowden, “Large Lexicon Detection of Sign Language Human Computer Interaction”, pages 88–97, 2007.
2. T. Cootes, G. Edwards, and C. Taylor, “Pattern Analysis and Machine Intelligence”, IEEE Transactions on, 23(6):681–685, 2001.
3. H. Hermansky, D. Ellis, and S. Sharma “Tandem Connectionist Feature Stream Extraction For Conventional HMM System ”, pages 1635–1638, 2000.
4. H. Hienz, B. Bauer, and K. Kraiss, “Hmm-Based Continuous Sign Language Recognition Using Stochastic Grammars”, Gesture-Based Communication in Human-Computer Interaction, pages 185–196, 1999.
5. T. Starner, J. Weaver, and A. Pentland, “Real-Time American Sign Language Recognition Using Desk And Wearable Computer Based Video”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(12):1371–1375, 1998.

AUTHOR(S) PROFILE

Revathi. P, received the M.Sc., degree in Computer Science from Thiruvalluvar University in 2016. She is a Research fellow, Department of Computer Science. Her research interest center around the performance analysis of computer systems, especially on cloud computing, optical networking data warehousing.