Portable Camera Based Assistive Pattern Recognition for Visually Challenged Persons

R. Savitha
M.E Communition Systems,
Department of ECE,
PREC,
Thanjavur , India.
Email-savitha.rajendran1985@gmail.com

R. Ramya
Assistant Professor,
Department of ECE,
PREC,
Thanjavur , India.
Email-ramya.ece14@gmail.com

***Abstract*—** Choosing clothes, food recognition and traffic signal analysis are major challenges for visually impaired persons. The existing automatic clothing pattern recognition is also a challenging research problem due to rotation, scaling, illumination, and especially large intra class pattern variations. This project, a camera based assistive framework is proposed to help blind persons for identification of food pattern, clothe pattern and colors in their daily lives. The existing traffic signal using sensors method is difficult to analysis and many components used. A camera based traffic signal analysis method easy to handle, to provide clear traffic signal analysis and reduce the time delay. The system contains the following major components 1) a camera for capturing clothe, food and traffic signal images, a microphone for speech command input; 2) data capture and analysis to perform command control, recognize clothe patterns, food patterns and traffic signal identification by using a wearable computer and 3) a speaker to provide the name of audio outputs of clothe patterns and colors, food patterns and traffic signal analysis, as well as system status. To handle the large intra class variations, a novel descriptor, Radon Signature is proposed to capture the global directionality of clothe patterns, food patterns and traffic signal analysis. To evaluate the effectiveness of the proposed approach CCNY clothes Pattern dataset is used. Our approach achieves 92.55% recognition to improve the life quality, do not depend others.

***Keywords*—** *Pattern recognition, CCNY, random signatures, statistical analysis .*

*****

I. INTRODUCTION

Visually challenged people worldwide, 45 million are blind. New development in computer revelation, digital cameras and portable computers make it sufficient to aid these individuals by rising camera based products that join computer vision machinery with other existing profitable goods such optical character recognition (OCR) systems.[1] There are three categories of steering systems i) vision improvement, ii) vision substitution and iii) vision replacement. Vision enhancement involves input from a camera, dispensation of the information, and production on a visual show. In its simplest shape it may be a small head-mounted camera by means of the output on a head-mounted diagram display. Vision substitution involves displaying the information openly to the image cortex of the human being brain or via the optic nerve. Vision replacement is similar to vision enrichment but with the productivity being non-visual – typically tactual or aural or some mixture of the two. ETAs fit in to the vision replacement systems. They change information about the surroundings that would usually be relay from side to side vision into a form that can be convey from side to side another sensory modality.

Development of a sensing device that can provide a sufficient perceptual substrate for persons with visual impairments to orient themselves and travel confidently has been a persistent. Designers of camera-based navigation devices may wish to consider the preferences of our sample by incorporating a query-based interface with simple keypad input and speech output, and to include in their object recognition efforts the goal of identifying architectural features that are significant to users who are blind in navigation.[2] Computer vision-based approach offer considerable compensation with deference to persons system and constitute a talented option to address these troubles. By the way of visual Simultaneous Localization and map techniques, it is possible in the direction of construct an incremental map of the situation, as long as at the same point in time the location and spatial orientation of the user within the environment. In addition, compared with other sensory modalities, computer dream can also give very rich and valuable insight in order of the surroundings such as obstruction discovery or 3D sight sympathetic A microphone, a computer, and a Bluetooth earpiece for audio explanation of clothing patterns and colors.

A camera placed upon a couple of sunglasses is used to take into custody clothing images. The clothing pattern and colors are described to blind users by a vocal display with negligible interruption to hearing. The system be able to be controlled by speech input from side to side a microphone.[7] In order to make easy blind users to interact, language commands effort from a microphone is old to provide function assortment and scheme control.

II. RELATED WORK

In [1] Eric Nowak, Frederic Jurie presents Sampling Strategies for Bag-of-Features Image Classification. The essential idea is to treat imagery as loose collection of autonomous patches, sample a envoy set of patch from the picture, evaluate a visual descriptor vector for each scrap separately, and by the resultant allocation of samples in descriptor gap as a description of the image. The four major accomplishment stages are thus how in the direction of replica patch, how to clarify them, how to typify the ensuing distributions and how to classify imagery base on the result. They attentiveness the primary issue, showing experimentally that for a delegate selection of usually old test database and for reasonable to large information of sample, random example gives equal or better classifiers than the complicated multi scale attention operators that be in ordinary use.

In [2] Xiaodong Yang, Shuai Yuan, and YingLi Tian presents Recognizing Clothes Patterns for Blind People by Confidence Margin based Feature Combination. a new method To categorize clothes pattern into 4 categories: strip, web, particular, and model less. While presented feel investigation methods primarily alert on texture unstable with characteristic prototype change, they cannot attain the similar level of correctness intended for clothes pattern credit because of the large intra-class variation in each clothing prototype category. To resolve this difficulty, we take out both structural characteristic and arithmetical characteristic starting image wavelet associate bands. In addition, we expand a new characteristic mixture system base on the poise margin of a classifier to merge the two types of features to form a story local image descriptor in a compact and discriminative layout.

In [3] Yingli Tian1, Xiaodong Yang1, and Aries Arditi presents Computer Vision-Based Door Detection for Accessibility of Unfamiliar Environments to Blind Persons. A generic arithmetical door copy is build to perceive door by combine limits and corner. Still, additional geometric in revolve is unemployed to decide door from further bits and pieces with related size and form. The forcefulness and generalizability of the projected detection algorithm are evaluate against a challenging file of doors composed from a variety of environment over a wide choice of colors, textures, occlusions, illuminations, scale, and views. There have been many efforts to study screen steering and way result with the final goal of rising useful voyage aids for shade community but very few have met with more than restricted success.

In [4] j. Zhang and m. Marsza_lek, s. Lazebnik presents local features and kernels for classification of texture and object categories: a comprehensive study. Evaluate The presentation of this advance with dissimilar key tip detectors and descriptors, as healthy as unlike kernel and classifiers. We next conduct a relative assessment with quite a few high-tech respect stages on four surface and five purpose database. On most of these database, our discharge exceed the preeminent report consequences and achieve equal performance on the rest. Finally, we examine the power of backdrop correlation on credit presentation via extensive tests on the PASCAL database, for which ground-truth thing localization in order is available.

In [5] Yong Xu1, Hui Ji presents Viewpoint invariant texture description using fractal Analysis. The MFS is invariant beneath the bi-Lipchitz chart, which include view-point changes and non-rigid deformations of the feel outside, as well as local affine light change. It provide an well-organized scaffold combine global spatial invariance and local robust capacity

III. PROJECT EFFORTS

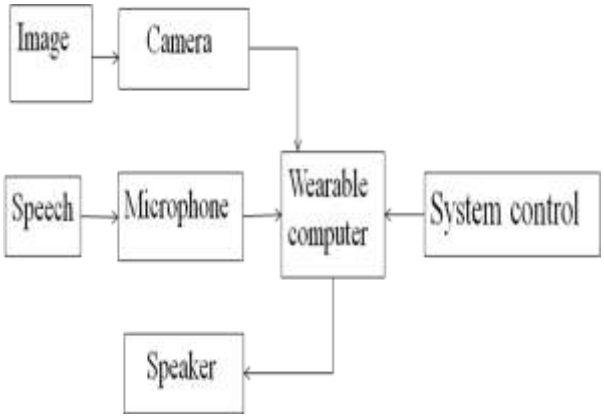

Fig 1: Overall Block Diagram

A portable camera can be mounted upon a pair of sunglasses or wrist watch is used to capture the food images or clothe images. These images are applied through the input of wearable computer. The speech commands input from a microphone, which is used to select the type such as food items or clothe patterns. The microphone output applied through a wearable computer.[8] The wearable computer to perform the command control, patterns recognize, and color identify. These wearable computer output applied through speaker input, the audio output provides the sound of food and color. Camera based system to help visually impaired people to recognize clothing patterns, food and colors. The system contains three major components.[3]

Fig 2: Clothing Pattern Recognition

1. Sensors including a camera for capturing clothes images, a microphone for speech control input and speakers for audio result.
2. Wearable computer perform command control, clothing pattern recognition, food recognition and color identification. Example for wearable computer mini-computer, Smartphone etc.
3. Audio outputs to provide clothing pattern recognition results and colors, food recognition results, as well as system status.

Fig 3: Wrist Computer

Controls inputs are switch on the system, low battery, turn off the system, system restart and stop function. Users to interact with the system, speech commands input from a microphone are used for provide function selection and system control. Blind user can verbally request the function clothing recognition aid to execute or food recognition or traffic signal analysis to execute. The recognition results will be presented to the blind user as audio outputs together with recognized, not recognized, and starts a new function. As for the recognized function, the next level functions include clothe pattern and colors /food patterns /traffic signal analysis to announce the recognized pattern.[9]

IV. PATTERN RECOGNITIONS

Pattern recognition Radon Signature, statistical descriptor (STA), and scale invariant feature transform (SIFT) are used.

A. Radon Signature

Radon Signature is based on the Radon transform which is generally used to detect the standard orientation of an image. The image is then rotated according to this prevailing direction to accomplish rotation invariance. The Radon transform of a 2-D function $f(x, y)$ is given as[8]

$$R(r,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\delta(r - x\cos\theta - y\sin\theta)\,dx\,dy \quad (1)$$

Where r is the perpendicular distance of a projection line to the initial point and θ is the angle of the projection line. The directionality of an image can be represent by Var (r, θi), the variances of r under a certain projection direction θi

$$\mathrm{var}(r,\theta_i) = \frac{1}{N}\sum_{j=0}^{N-1}(R(r_j,\theta_i) - \mu(r,\theta_i))^2 \quad (2)$$

$$\mu(r,\theta_i) = \frac{1}{N}\sum_{j=0}^{N-1} R(r_j,\theta_i) \quad (3)$$

where R(rj, θi) is the projection value at perpendicular distance of rj and projection direction of θi. μ (r, θi) is the expected value of R(r, θi) N is the number of sampling bins in each projection line.

B. Statistical descriptor (STA)

Statistical features are well adapted to analyze textures which lack background clutter and have uniform statistical properties. Discrete wavelet transform provides a generalization of a multire solution spectral analysis tool. Four statistical values of an image can be calculated including variance, energy, uniformity, and entropy to all wavelet sub bands.[4]

$$\text{variance} = \sum_{i=0}^{L-1}(z_i - m)^2 p(z_i)/(L-1x)^2$$

$$\text{energy} = \sum_{i=0}^{L-1}(z_i - m)^3 p(z_i)/(L-1x)^2$$

$$\text{uniformity} = \sum_{i=0}^{L-1} p^2(z_i)$$

$$\text{entropy} = -\sum_{i=0}^{L-1} p(z_i)\, log_2\, p(z_i)$$

Where z_i and $p(z_i)$, $i = 0, 1, 2, \ldots, L-1$ is the intensity level and corresponding histogram. L is the number of intensity levels. $m = L-1\ i=0\ z_i p(z_i)$ is the average intensity level.

C. Scale Invariant Feature Transform

SIFT descriptor as the illustration of interest points based on the following reasons: the descriptor with 128 dimensions is compact and fairly distinct; the representation with careful design is forceful to variations in illumination and viewpoints; an extensive comparison against other local

 ISSN: 2321-8169
1030 - 1034

image descriptors observed that the SIFT descriptor performed well in the context of image matching.

The bag of- words (BOW) method is further applied to collective extracted SIFT descriptors by labeling each SIFT descriptor as an illustration word and counting frequencies of each illustration world.

D. UIUC

1. The balancing relations linking the planned total and local facet channels

2. The advantage of our future technique more than the state-of-the-art consistency categorization approach in the background of clothes prototype recognition[8]

3. The sweeping statement of our move toward on the customary texture classification.

The UIUC Texture dataset which is a entrenched customary texture dataset. It contain 1000 calibrate and unregistered imagery through the promise of 640 × 480.[10] There are 25 texture classes by means of 40 images for each class. The feel images present turning round, scale and view change, and non inflexible facade bend.

V. FLOW CHART

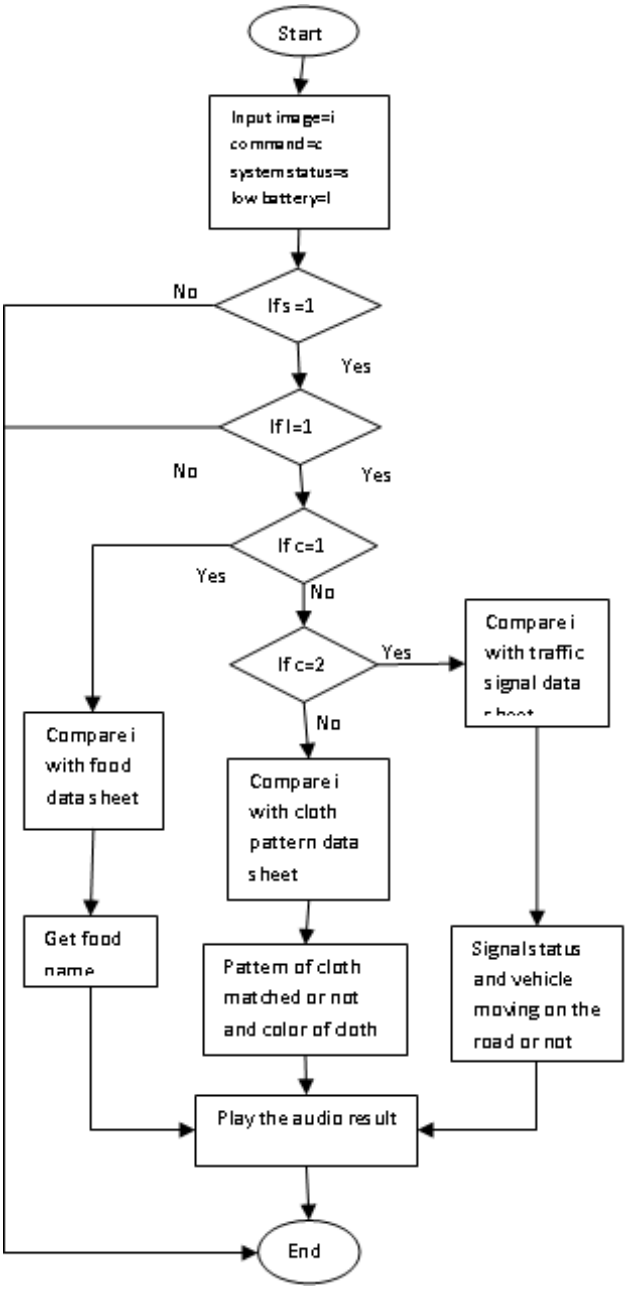

Fig 4: Flow Chart

Start the program, let we consider the input image is i, and speech command is consider c , and the system status for on or off condition is consider s , and low battery level is consider l. If the system is on s=1, if the condition is yes go to next stage and the system is off the condition s is not equal to 1 go to end. If the low battery level is l=1 go to end and l is higher level go to next stage.[3] If the command is c=1 go to food identification images are compared, and get a food name to audio output. If the command c is not equal to 1 go to next condition. If c=2 go to traffic signal analysis , the traffic signal images compared, and get a traffic signal audio output. If c is not equal to 2 go to clothe patterns, compare clothe images to provide a audio output of clothe patterns and colors.

VI. EXPECTED RESULT

The comparisons of different feature channels and their combinations validate our intuition that the effectiveness and complementarities of our proposed feature channels. The detailed recognition accuracies. The percentages of training images per class are 10%, 30%, 50%, and 70%, respectively. The recognition accuracy of SIFT+STA+RadonSig using 30% of the images as the training set is comparable or even better than that of other feature channels using 70% of the images as the training set. This observation demonstrates another merit of our proposed approach that it is able to achieve a desirable result by using much less training data.

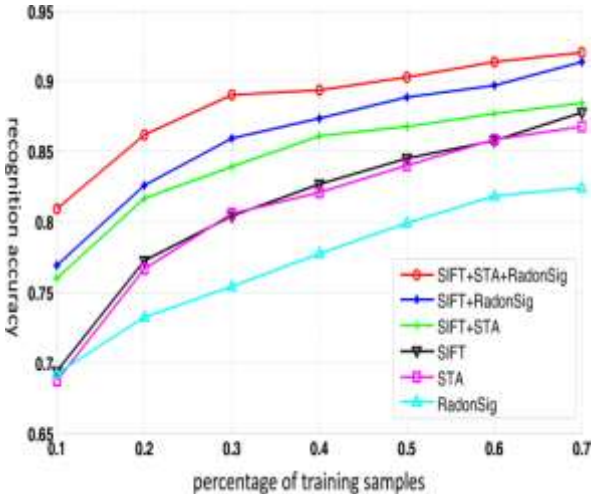

Fig 5 : Comparative evaluation on recognition accuracies

CONCLUSION

A system to recognize clothing patterns and colors, food & traffic vehicle to help visually impaired people in their daily life. Experimental results demonstrate that our proposed method significantly outperforms the state-of-the-art methods in clothing pattern recognition. Furthermore, the performance evaluation on traditional texture datasets validates the generalization of our method to traditional texture analysis and classification tasks.[8] The blind users expressed a desire for faster speech feedback in order to gain more information. Some blind participants expressed a desire to have the camera on a cap instead of on glasses and to have the function available on mobile phones. In this systems are being developed to improve the life quality and safety for those with special needs including indoor navigation and way finding, display reading, banknote recognition, rehabilitation, etc. built a clothing recommendation system for specific occasions.[2] With real-time updates, blind users will be able to better use spatial memory to understand the surrounding environment**.** The method also provides new functions to improve the life quality for blind and visually impaired people. The Same system can be used for more than one application.

Easy to handle the system. Blind people do not depend other. It is Portable one. Improve the life quality of blind and visually impaired people.

REFERENCES

[1] A.Arditi and Y.Tian, “User interface preferences in the design of a camera based navigation and way finding aid,” J. Visual Impairment Blindness, vol. 107, no. 2, pp. 18–129, 2013.

[2] D.Dakopoulos and N. G. Bourbakis, “Wearable obstacle avoidance electronic travel aids for the blind: A survey,” IEEE Trans. Syst., Man, Cybern. C, vol. 40, no. 1, pp. 25–35, Jan. 2010.

[3] L. Davis, S. Johns, and J. Aggarwal, “Texture analysis using generalized co-occurrence matrices,” IEEE Trans. Pattern Anal. Mach. Intel., vol. PAMI-1, no. 3, pp. 251–259, Jul. 1979.

[4] D. Gould, “The making of a pattern,” Vogue Patterns, 1996.

[5] R. Haralick, “Statistical and structural approaches to texture,” Proc. IEEE, vol. 67, no. 5, pp. 786–804, May 1979.

[6] F.Hasanuzzaman, X.Yang, and Y. Tian, “Robust and effective component based banknote recognition for the blind,” IEEE Trans. Syst., Man, Cybern. C, vol. 42, no. 6, pp. 1021–1030, Nov. 2012.

[7] A. Huete, J. Victores, S. Martinez, A. Gimenez, and C. Balaguer, “Personal autonomy rehabilitation in home environment by a portable assistive robot,” IEEE Trans. Syst., Man, Cybern. C, vol. 42, no. 4, pp. 561–570, Jul. 2012.

[8] K. Khouzani and H. Zaden, “Radon transform orientation estimation for rotation invariant texture analysis,” IEEE Trans. Pattern Anal. Mach. In tell. vol. 27, no. 6, pp. 1004–1008, Jun. 2005.

[9] I. Kocur, R. Parajasegaram, and G. Pokharel, “Global data on visual impairment in the year 2002,” Bulletin World Health Org., 2004.

[10] S. Lam, “Texture feature extraction using gray level gradient based on co-occurrence matrices,” in Proc. Int. Conf. Syst., Man Cybern., 1996, pp. 267–271.