International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 05 | May-2016 www.irjet.net p-ISSN: 2395-0072

INDOOR GUIDANCE FOR VISUALLY IMPAIRED USING COMPRESSIVE SENSING

Ferdinand.J[1], T.C.Subbulakshmi[2],

[1]PG Scholar, Dept. Of IT, Francis Xavier Engineering College,Tirunelveli
[2]Dept. Of IT, Francis Xavier Engineering College,Tirunelveli
ferdin27@gmail.com

---***---

Abstract - *For assisting blind people to recognize their indoor objects, a new portable camera based method is introduced here. Coarse description is used here to perform the recognition task. With this technique, the presence and absence of different objects in a given query image is described. Regardless of their position within the query image it conveys the list of most likely objects present in the query image. It controls the processing time by sacrificing irrelevant information details and expands the recognition task to multiple objects. With this technique, the perception of the blind people to his direct contextual environment can be increased. The two image multilabeling strategies which have different similarity computation is used to address the coarse description issue. First, the Euclidean distance measure is used and second one uses Gaussian process (GP) estimation which relies on semantic similarity measure. The idea is to compare the given query image with the entire set of training images that are stored offline. These offline training images are available with their associated binary descriptors. Compressive sensing provides a compact image representation to achieve fast computation capability. The two methods Euclidean Distance measure and Semantic Similarity is compared and the efficiency is studied.*

Key Words: Object recognition, Euclidean measure, Gaussian process Compressive Sensing and Digital Image Processing.

1.INTRODUCTION

The cause of this work emerged from the need of blind men and women to advisor themselves within building. It is most important that the blind and the visually impaired have their independence, that's, discontinue relying on different folks to help them perform day-to-day pursuits. Happily, there are the guide dogs and guide canes that help them to walk through the streets. Nonetheless, there's a few things finished to help them to move within buildings safely and with no need to depend upon different men and women. [1- 2]

There are a number of approaches devoted to scene recognition which have been proven to be notably victorious in recognizing outside scenes. Nonetheless, when these methods are tested on indoor scene classes the results drop dramatically for most customary indoor scenes. There are two essential explanations for the slow growth in this area [3].

The primary purpose is the shortage of a enormous test bed of indoor scenes where to train and test one of a kind procedures. The second reason is that with a purpose to improve indoor scene recognition performance we have to advance image representations certainly tailor-made for this mission. The primary concern is that even as most out of doors scenes can also be well characterized by global image properties this is not real of all indoor scenes [4]. Some indoor scenes (e.G. Corridors) can indeed be characterized by using global spatial properties but others (e.G bookstores) are better characterized by using the objects they include. For most indoor scenes there's a vast variety of both local and global discriminative information that desires to be leveraged to solve the task of recognition.

1.1 Indoor Scene Description

Figure.1 depicts the proposed image multi labeling process. The underlying perception as hinted previous is to evaluate the viewed query image with an complete set of training snap shots. The binary descriptors of the k most identical images are considered for successive fusion in order to multi label the given query picture. This fusion step, which objectives at reaching better robustness within the decision system, is situated on the easy majority-established vote utilized on the k most an identical photos (i.e., an object is detected within the question picture provided that, amongst the k training pictures, it exists once for k=1, at least twice for k=3, and at least thrice for okay=5). For that reason, each training image within the library earns its possess binary multi labeling vector (or conveniently snapshot descriptor), which feeds the fusion operator.

The events for establishing such vector for a given training image is to visually examine the existence of each object inside a predefined list in the photograph. If an object exists inside a given depth variety forward assessed with the aid of visible inspection of the regarded training image (e.g., 4 meters) then a ‘1’ is assigned to its related bin in the vector, or else a ‘0’ value is retained.

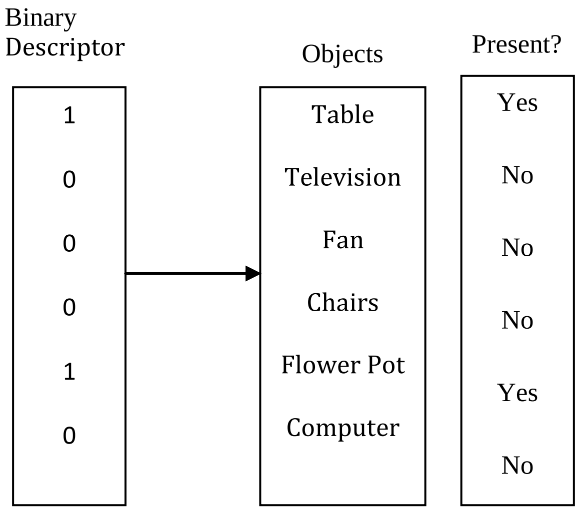

Figure 1: Relationship between the binary descriptor and the predefined set of objects.

In scene classification our intention is to be taught a mapping from pics x to scene labels y. For simplicity, in this section we assume a binary classification surroundings. That is, each and every yi 2 (1,−1) is a binary label indicating whether an photograph belongs to a given scene class or now not. To model the multiclass case we use the standard strategy of training one versus all classifiers for each scene; at experiment, we predict the scene label for which the corresponding classifier is most positive. Nevertheless, we wish to word that our model can be simply adapted to an explicit multiclass training method. The underlying perception as hinted prior is to compare the viewed query image with an whole set of training pics.

2. COMPRESSIVE SENSING

In Fig.2 compressive sensing is a brand new type of sampling theory, It assures the reconstruction of sparse signals and images from the less or incomplete information. The traditional methods that are used for reconstruction of images needs its sampling rate to be twice the highest frequency. It also states that the total number of measurements for the discrete signal must be greater than or equal to its original length to make sure the reconstruction of the image. But the Compressive sensing states that with few measurements about the signal it is possible to reconstruct the original signal when assumptions or prior knowledge of the signal is available. It takes benefit of redundancy in the signal. Sparsity of the signal plays a major role in compressive sensing. It minimizes the number of number of non zero elements in the signal.

Compressive sensing is a way to achieve a sparse representation of a signal. It is predicated on the idea to take advantage of redundancy (if any) in the signal. Signals like images are sparse, as they comprise, in some representation domain, many coefficients close to or equal to zero. The fundamental of the CS idea is the potential to recover with relatively few measurements [5].

$$y = \varphi x \quad (1)$$

φ – measurement matrix, x – original image, y – measurement results.

The Three stages of compressive sensing are, sparse representation, measurement matrix, and signal reconstruction.

Compressive sensing is useful when,

1. Signals are sparse in nature.
2. Measurement at transmitting end costs much.
3. Receiver end computation is cheap.

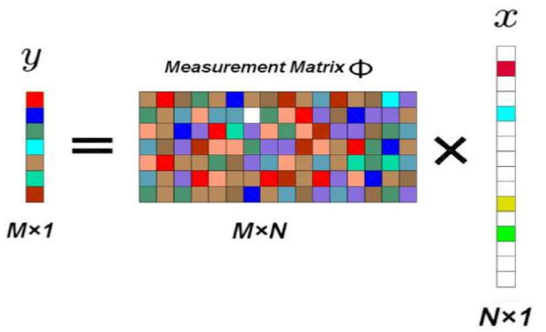

Figure 2: Compressive Sensing sampling process

3. PRINCIPAL COMPONENT ANALYSIS

Principal Component Analysis (PCA) is a statistical system that uses an orthogonal transformation to transform a collection of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of components is not up to or equal to the number of original variables. This transformation is outlined in this kind of approach that the principal component has the most important viable variance, and each and every succeeding aspect in flip has the perfect variance possible under the constraint that it's orthogonal to the preceding components. The ensuing vectors are an uncorrelated orthogonal foundation set. The principal components are orthogonal given that they are the eigenvectors of the covariance matrix, which is symmetric. PCA is sensitive to the relative scaling of the original variables.

PCA is the simplest of the actual eigenvector-established multivariate analyses. By and large, its operation may also be thought of as revealing the interior structure of the data in a way that best explains the variance within data. If a multivariate dataset is visualised as a suite of coordinates in a high-dimensional data space , PCA can provide the person with a lower dimensional picture, a projection or "shadow" of this object when viewed from its most informative standpoint. This is performed by using first principal components in order that the dimensionality of the converted data is reduced.

PCA is intently involving factor analysis. Factor analysis almost always accommodates more area particular assumptions in regards to the underlying structure and solves eigenvectors of a rather distinctive matrix. PCA can also be involving canonical correlation analysis (CCA). CCA defines coordinate systems that optimally describe the cross-covariance between two datasets while PCA defines a brand new orthogonal coordinate approach that optimally describes variance in a single dataset.

The transformation T = X W maps an data vector x(i) from an long-established space of p variables to a brand new space of p variables that are uncorrelated over the dataset. However, not all the principal components need to be kept. Preserving most effective the first L principal components, produced by using primary L loading vectors, gives the truncated transformation,

$$T_L = XW_L$$

where the matrix TL now has n rows but handiest L columns. In different phrases, PCA learns a linear transformation t=W^T x,x∈ R^p,t∈R^(L)where the columns of p × L matrixW form an orthogonal basis for the L features (the components of representation t) which are decorrelated. By development, of all of the transformed matrices with only Lcolumns, this score matrix maximises the variance in the original data that has been preserved, even as minimising the whole squared reconstruction error $||TW^T - T_L W_L^T||_2^2$ or .

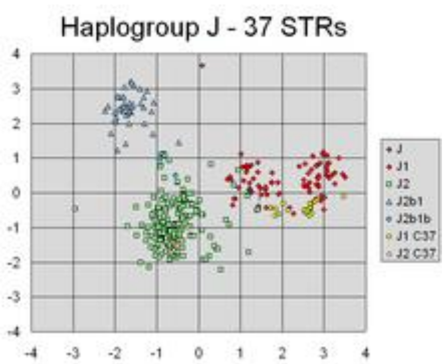

Figure 3: PCA Analysis Scatter plot

In Fig.3, principal component analysis scatter plot of Y-STR calculated from repeat-count values fourty seven Y-chromosomal STR markers from three hundred and fifty four individuals. PCA has efficiently discovered linear combos of the different markers that separate out one-of-a-kind clusters similar to distinctive lines of individuals' Y-chromosomal genetic descent.

Such dimensionality reduction is usually a very valuable step for visualizing and processing high-dimensional datasets, while nonetheless keeping as so much of the variance within the dataset as possible. For instance, opting for L = 2 and maintaining most effective the primary two most important components finds the 2-dimensional plane via the high-dimensional dataset where the data is most spread out, so if the data involves clusters these too could also be most spread out, and as a consequence most seen to be plotted out in a two-dimensional diagram; whereas if two directions via the data (or two of the normal variables) are chosen at random, the clusters could also be so much less spread apart from every different, and could actually be more likely to radically overlay each and every other, making them indistinguishable.

In a similar fashion, in regression analysis, the higher the number of explanatory variables allowed, the bigger is the threat of over fitting the mannequin, producing conclusions that fail to generalise to other datasets. One method, in particular when there are robust correlations between different possible explanatory variables, is to diminish them to a few principal components after which run the regression in opposition to them, a process referred to as principal component regression.

Dimensionality reduction will also be right when the variables in a dataset are noisy. If every column of the dataset comprises independent Gaussian noise, after which the columns of T will also incorporate in a similar way identically disbursed Gaussian noise. Nevertheless, with more of the whole variance concentrated in the first few components in comparison with the same noise variance, the proportionate result of the noise is much less—the primary few components achieve a larger sign-to-noise ratio. PCA therefore can have the result of concentrating a lot of the signal into the first few components, which will usefully be captured with the aid of dimensionality reduction; even as the later principal components is also dominated by means of noise, and disposed of without much loss.

4. SEMANTIC SIMILARITY

Given two pictures I_1 and I_2 along with their corresponding binary descriptors b_1 and b_2. I define the quantity SS_{I_1,I_2} as the semantic similarity between I_1 and I_2. In targeted, this measure express the ratio inclusion of I_2 in I_1, that is the number of objects of I_2 (nonetheless represented as ones in b_2) present additionally in I_1 (i.e., still represented as ones in b_1). For this reason, the better the SS_{I_1,I_2} the (sematically) closer I_2 to I_1. Mathematically, it is expressed by:

$$SS_{I_1,I_2} = \frac{\sum_{i=1}^{N} b_1(i) . b_2(i)}{\sum_{i=1}^{N} b_1(i)} \quad (2)$$

The multilabelling process that is based on the sematic similarity prediction is articulated over two phases:

Training Phase: First, compute the SS values between all couples of training pictures. Then, train as many GP regressors as the quantity of training pictures. Each GP regressor will be learned to foretell SS_{I_p,I_i}, that is the semantic similarity between a given image I and the training picture IP to which the GP regressor is associated. The supervised training of the p-th predictor is performed through giving: i) in input the CS coefficients corresponding to each and every training image I_i; and ii) in output as target the SS_{I_p,I_i} values.

Operational Phase: Feed each GP predictor with CS coefficient vector of the query picture I to estimate all SS_{I_p,I_i} values, i.e., the similarity between I and each training pictures I_p. Subsequently, pick up the k binary descriptors associated with the training pictures corresponding to the k best values for successive fusion, and infer the multilabeling of the query picture.

Gaussian Process Regression: Consistent with the GP formulation, the learning of a machine is expressed in terms of a Bayesian estimation problem, where the parameters of the machine are assumed to be random variables which can be a-priori jointly drawn from a Gaussian distribution. In larger detail, let us consider $X = \{X_i\}_{i=1}^{N}$ a matrix of input data representing our N training images and where $x_i \in \Re^{N_c}$ represents a vector of processed features, Specifically the N_c CS coefficients associated with the i-th training image.

5. RESULTS

The images adopted for evaluating the efficiency of the proposed method refers to two different buildings (Shown in Figures 4 to 7). It is noteworthy that the training images for both datasets were selected to cover all the predefined objects within the considered indoor environment.

As noted above, a record of objects of interest must be predefined. Thereupon, we have selected the objects deemed to be the foremost important ones in the regarded indoor environments. Concerning the first dataset, 10 objects were selected as follows: 'Table', 'Fan', 'People', 'Laptop', 'Chairs', 'Jar', 'Sofa', and 'Elevator'. For the second set, the list was the following: 'Steps', 'Heater', 'Hall', 'Board', 'Rooms', 'Container', 'Table', 'People', 'Pillar'. Given that the results got from the first dataset, it comes out that, in overall phrases, each the semantic similarity-established compressed sensing (SSCS) and Euclidean distance-based compressed sensing (EDCS) approaches perform equivalently on an normal over the SEN and SPE accuracies. However, the SSCS process yields a greater sensitivity at the same time the EDCS shows a better specificity. As for the 2d dataset, the EDCS performs somewhat higher through taking the averages for k=1. For the other okay values, the SSCS outperforms. This is explained by means of the fact that the EDCS depends on measuring the similarity of the CS coefficients, yet measuring the obvious similarity between the pics, which is more likely to assurance the query picture surely resembles to the primary closest photograph from the library (for k=1).

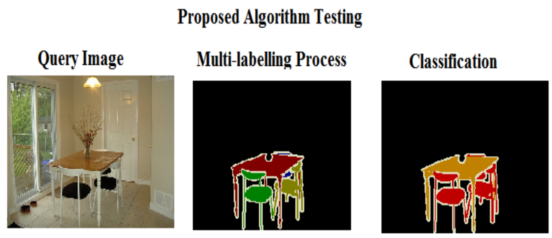

Figure 4: Classified Images for the subset of scene categories

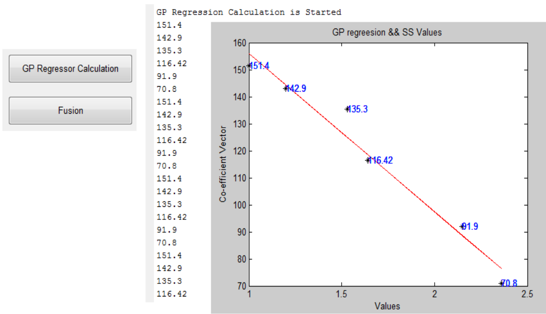

Figure 5: GP Regression Calculation

In line with the GP method, the learning of a machine is expressed in terms of a Bayesian estimation problem, where the parameters of the machine are assumed to be random variables which can be a priori collectively drawn from a Gaussian distribution. It represents the capacity of the model to fit the info shown in above figure.

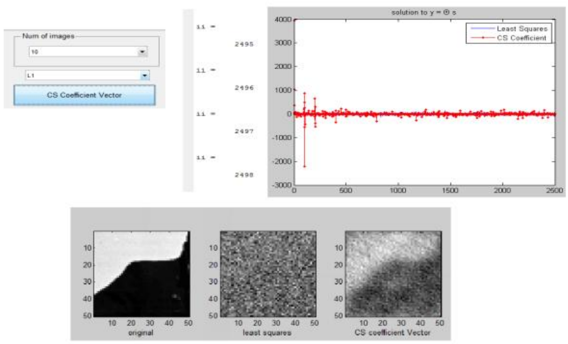

Figure 6: CS Co-efficient Vector Calculation

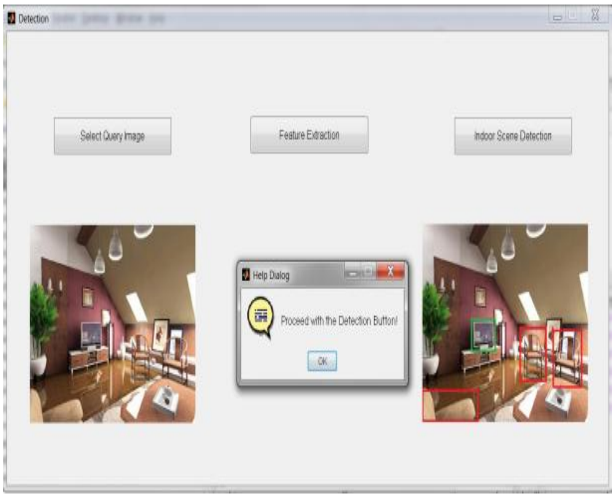

Figure 7: Indoor Scene Detection

Results of SIFT-based method achieved on the above figures. Detected objects are: Chairs, Television. The chairs and television label is set as 1 in the binary descriptors. To differentiate the labeling concept the chairs label is draw as red rectangle and another object such as television is drowned as green rectangle.

3. CONCLUSIONS

A original way of tackling the challenge of object realization underneath a blind rehabilitation prospect is the reliance on detecting one targeted type of objects. As to relay extra expertise on the scene beneath analysis, broadening the emphasis to a couple of objects becomes quintessential however raises implementation issues as a result of the very tight time constraint. To this finish, this paper introduces a novel multi-object detection strategy for indoor scenes by means of coarse picture description, which is fulfilled by using multilabeling an photo obtained by means of a camera installed on the consumer. Coarse photograph description was regarded with the goal to increase the notion and the comprehension of a blind man or woman to his/her nearby objects in an indoor atmosphere. The inspiration is to utilize an offline ready library inclusive of distinct photos captured from unique points dispensed all over the considered indoor atmosphere. Picture illustration was once handled via a compressive sensing-centered procedure in an effort to guarantee compactness and as a consequence brief picture analysis time. The query photo is coarsely described via fusing a given quantity of most an identical pix from the library. In this context, the similarity proposal was once implemented by means of proposing two methods, a common Euclidean distance method and a semantic-centered similarity. The latter one is favored over the former one for the reason that it gauges the resemblance between graphics on the basis of their semantic content and no longer on their spectral appearances.

REFERENCES

[1] Ulrich, J. Borenstein, "The GuideCane-applying mobile robot technologies to assist the visually impaired," IEEE Trans. Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 31, no. 2, pp.131-136, 2001.

[2] S. Shraga, J. Borenstein, Y. Koren, "Auditory guidance with the navbelt-a computerized travel aid for the blind," IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 28. no. 3, pp. 459-467, 1998.

[3] M. R. Strakowski, B. B. Kosmowski, R. Kowalik, P. Wierzba, "An ultrasonic obstacle detector based on phase beamforming principles," IEEE Sensors Journal, vol. 6, no.1, pp. 179-186, 2006.

[4] S. Pundlik, M. Tomasi, G. Luo, "Collision Detection for Visually Impaired from a Body-Mounted Camera," IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 41-47, 2013.

[5] V. Pradeep, G. Medioni, J. Weiland, "Robot vision for the visually impaired," IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 15-22, 2010.