IJIRR

International Journal of Information Research and Review
Vol. 04, Issue, 02, pp.3729-3734, February, 2017

Research Article

BLUR DETECTION AND EXTRACTION OF BACKGROUND IMAGES IN MOBILE DISPLAY DEVICES USING CURRENT TECHNOLOGIES

***Srilekha, T.**

Department of Computer Science and Engineering, VSBCETC, Coimbatore 642109

ARTICLE INFO

Article History:

Received 17th November, 2016
Received in revised form
15th December, 2016
Accepted 19th January, 2017
Published online February, 28th 2017

Keywords:

Manifold Algorithm,
Content Based Image Retrieval,
Keyword Propagation,
Active Learning,
Offline training,
Convergence.

ABSTRACT

Images of high (or) super resolution will be good to see. And having excellent clarity in all kinds of devices. Nowadays all mobile displays are not providing the users expecting clarity. Snaps taken from the mobile will be blurred (or) not clear sometimes. Reason for these problems are not using clarity lenses, shaking of mobile while taking snaps. For getting clarity in images first we should consider the pixels of the images. Pixels of an image should be very minute. If the image having very less spilt up in their pixels, then that image will look like blurred. So separation of pixels in images plays an important role here. For considering the pixels we should first use manifold algorithm. It split the images as pixels with good clarity and split as according to their texture also. In contrast to existing methods which train a binary classifier for each keyword, our keyword model is constructed in a straightforward manner by exploring the relationship among all images in the feature space in the learning stage. In relevance feedback, the feedback information can be naturally incorporated to refine the retrieval result by additional propagation processes. In order to speed up the convergence of the query concept, we adopt two active learning schemes to select images during relevance feedback. Furthermore, by means of keyword model update, the system can be self-improved constantly.

INTRODUCTION

Content-Based Image Retrieval (CBIR) is a long standing research problem in computer vision and information retrieval. Most of previous image retrieval techniques build on the assumption that the image space is Euclidean. However, in many cases, the image space might be a non-linear sub-manifold which is embedded in the ambient space. The former is not always consistent with human perception while the latter is what image retrieval system desires to have. Specifically, if two images are semantically similar, then they are close to each other in semantic space. In this paper, our approach is to recover semantic structures hidden in the image feature space such as color, texture, etc. In recent years, manifold learning has received lots of attention and been applied to face recognition, graphics, document representation, etc. These research efforts show that manifold structure is more powerful than Euclidean structure for data representation, even though there is no convincing evidence that such manifold structure is accurately present. Based on the assumption that the images reside on a low-dimensional sub manifold, a geometrically motivated relevance feedback scheme is proposed for image ranking, which is naturally conducted only on the image manifold in question rather than the total ambient space.

****Corresponding author: Srilekha, T.,***
Department of Computer Science and Engineering, VSBCETC, Coimbatore 642109.

Relevance feedback on image manifold

In many cases, images may be visualized as points drawn on a low-dimensional manifold embedded in a high-dimensional Euclidean space. In this paper, our objective is to discover the image manifold by a locality-preserving mapping for image retrieval. The Algorithm Let Ω denote the image database and R denote the set of query images and relevant images provided by the user. Our algorithm can be described as follows:

- Candidate generation. For each image xi ∈ R, we find its k nearest neighbors Ci = {y1, y2, …, yk}, yj∈Ω (those images in R are excluded from selection). Let C = C1 ∪C2 ∪…∪C|R|. We call C candidate image set. Note that R ∩ C = □.
- Construct sub graph. Construct a graph G(V), where V=R∪C. where ε is a suitable constant. The choice of ε reflects our definition of locality. We put an edge between xi and xj if dist (xi, xj) ≠ ∞.
- Distance measure on image manifold. To model the geodesic distances between all pairs of image points on the image manifold M, we find the shortest-path distances in the graph G. The length of a path in G is defined to be the sum of link weights along that path.
- Retrieval based on geodesic distance. To retrieve the images most similar to the query, we simply sort them

according to their geodesic distances to the query. The top N images are presented to the user.

- Update query example set. Add the relevant images provided by the user into R. Go back to step 1 until the user is satisfied.

Using Manifold Structure for Image Representation

In the previous section, we have described an algorithm to retrieve the user desired images by modeling the underlying geometrical structure of the image manifold. One problem of this algorithm is that, if the number of sample images is very small, then it is difficult to recover the image manifold. In this case, we propose a long-term learning approach to discover the true topology of the image manifold using user interactions. To be specific, we aim at mapping each image into a semantic space in which the distances between the images are consistent with human perception. The problem we are going to solve can be simply stated below:

Our proposed solution consists of three steps:

- Inferring a semantic matrix $B_{m\times m}$ from user interactions, whose entries are the distances between pairs of images in semantic space T . m is the number of images in database.
- The user provided information is incorporated into the LE semantic space. Note that, the LE semantic space is only defined on the image database. In other words, for a new image outside the database, it is unclear how to evaluate its coordinates in the LE semantic space.
- Given m pair vectors, (x_i , z_i) $(i = 1,2,\ldots,m)$, where x_i is the image representation in low-level feature space, and z_i is the image representation in LE semantic space, train a radial basis function (RBF) neural network f that accurately predicts future z value given x. Hence f(x) is a semantic representation of x. The space obtained by f is called RBFNN semantic space. Note that, $f(x_i) \approx z_i$. That is, RBFNN semantic space is an approximation of the LE semantic space. However, RBFNN semantic space is defined everywhere. That is, for any image (either inside or outside the database), its semantic representation can be obtained from the mapping function.

Experimental Results

In this paper, we focus on image retrieval based on user's relevance feedback to improve the system's short-term and long-term performances. The user can submit a query image either inside or outside the database. The system first computes low-level features of the query image and then maps it into semantic space using the learned mapping function. The system retrieves and ranks the images in the database. Then, the user provides his judgment of the relevance of retrieval. With the user's relevance feedback, the system refines the search result iteratively until the user is satisfied. The accumulated relevance feedbacks are used to construct and update the semantic space. We performed several experiments to evaluate the effectiveness of our proposed approaches over a large image dataset. The image dataset we use consists of 3,000 images of 30 semantic categories from the Corel dataset. Each semantic category contains 100 images. The 3,000 images are divided into two subsets. The first subset consists of 2,700 images, and each semantic category contains 90 images. The second subset consists of 300 images, and each semantic category contains 10 images. The first subset is used as training set for learning the optimal mapping function. The second subset is for evaluating the generalization capability of our learning framework. Three types of color features (color histogram, color moment, color coherence) and three types of texture features (tamura coarseness histogram, tamura directionary, pyramid wavelet texture) are used in our system. The combined feature vector is 435-dimensional. For each query, a short-term learning process is performed and the feedbacks are used to construct the semantic space. The retrieval accuracy is defined as follows: N relevant images retrieved in top N returns Accuracy = Four experiments are conducted. The experiment with the new retrieval algorithm on image manifold is discussed.

Retrieval on Image Manifold

We compare the performance of our proposed retrieval algorithm on image manifold with the relevance feedback approach described in Rui. We didn't compare it to other image retrieval methods because our primary purpose is to analyze the geometrical structure of the image space.

MATERIALS AND METHODS

Mobile User Guided Adaptation System

To guarantee personalized media consumption with best possible perceptual experience in user-centric multimedia applications, both mobile device access environments and mobile user perceptual experiences are properly taken into consideration in this proposed scheme. The mobile environments include mobile device capabilities and mobile user interfaces, while mobile user perceptual experiences are highly affected by the semantics of media, user individual preference, and presentation of the adaptation results.

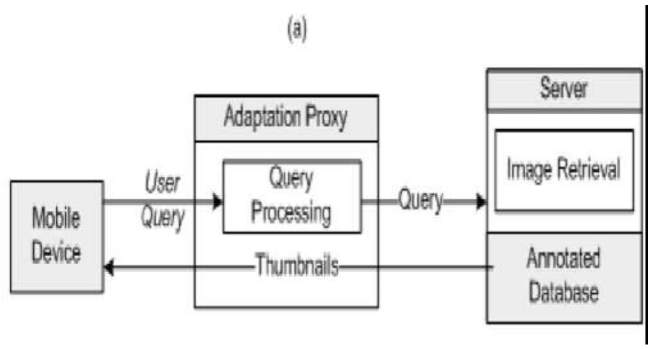


Fig 1. Query Processing

System Components

As shown in Fig. 1(a), the proposed system consists of 1) an adaptation proxy to process user request and feedback as well as to carry out semantic extraction, user preference learning, and adaptation; and 2) a server/database hosting original consumer photo content. We assume the annotation of the server side media content is processed offline while the user request and feedback processing is carried out in real time. .

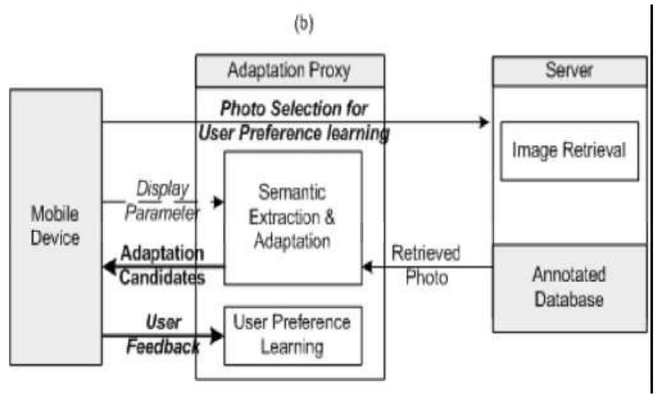


Fig 2. Semantic Extraction and Adaptation Process

System Workflow

The proposed system works in the following manner. These semantic keywords represent the key semantic concept for the desired media content for retrieval and can be used to match the associated annotations representing the semantics of the media content in the database.

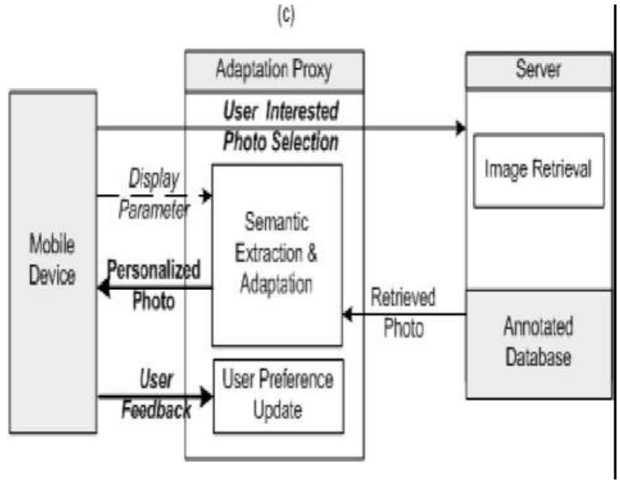


Fig 3. Semantic Extraction and Adaption

User guided semantic extraction

Human tends to view and understand images in terms of people and objects associated with underlying concepts in the real world. Hence, semantic analysis is an indispensable step towards extracting semantically important objects and learning user preferences for proper content selection in image adaptation to improve mobile users' perceptual experiences. Semantic gap is still a big challenge in computer vision. Fortunately, in our adaptation scenario, we do not need to perform a full semantic analysis for images, because for a given event, users tend to be interested in only a few objects. Hence, instead of carrying out full image semantic analysis, we extract key objects that users might desire. The key cue we can utilize to narrow down the semantic gap in our application is the user input queries. Although the limited mobile user interface usually does not allow very complicated input as query, the compact keywords provide simple yet useful information about the mobile users' intention. Although sometimes there is a departure between the mobile users' real intention and their query specification, the query is still informative enough to be utilized to extract semantically important objects as user preference relevant object candidates based on concept ontology

Bottom-Up Low Level Feature Extraction

In the bottom-up approach, salient regions are generated based on low level features. First, a raw saliency map is calculated based on low level features such as intensity and color. Finally, we represent the original image with regions of different saliency values by averaging the saliency within each region R and use the average value as the initial probability that the region belongs to the desired object.

User Guided Top-Down High Level Semantic Extraction

In the top-down approach, we develop the event specific classifier to obtain the high level features of the user interested objects with different semantic importance in the given events. The probability of the SIFT word appears on the object and the non-object area, respectively; and, the conditional probability of the color bin on the object and non-object region, respectively. The training is carried out offline for each event and these probabilities will be used as conditional probability in the following Bayesian fusion module.

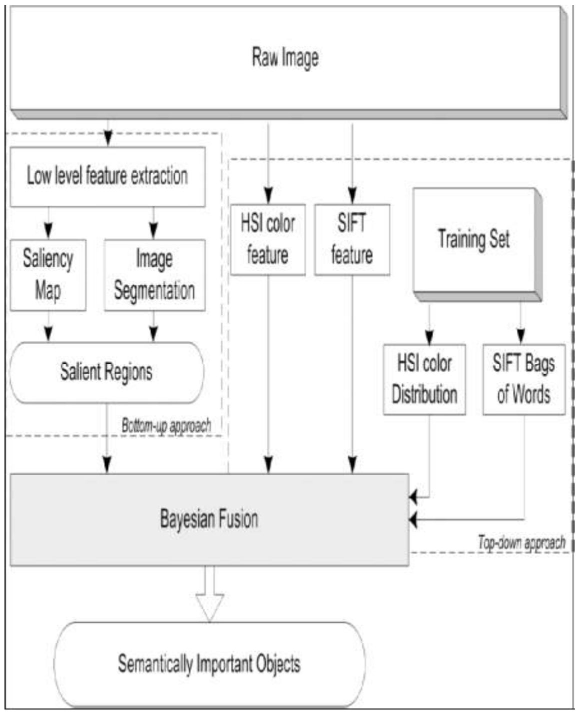


Fig 4. Bayesian Fusion

Key Objects Localization by Bayesian Fusion

After obtaining the results of bottom-up salient regions and the top-down high level features of main objects using event specific classifier, we fuse them to find the semantically important objects that will likely match with the user's interests. The fusion is designed via Bayesian principle to obtain the posterior belief map of a class of semantic objects. For each region, we consider the bottom-up salient regions as a priori of the region belonging to the class of semantic objects.

User Preference Learning For Adaptation

As discussed earlier, due to the limited mobile device user interfaces, it is usually not allowed to provide complicated inputs for mobile users to describe their desired content details very accurately.

Moreover, because of mobile users' different background, they tend to have different interests in concepts even if they input the same query. In CBIR systems, to bridge the intention gap in retrieving more relevant images consistent with user's interests, relevance feedback techniques have been developed to capture the subjectivity of human perception of visual content by dynamically updating weights of different features based on the user's feedback.

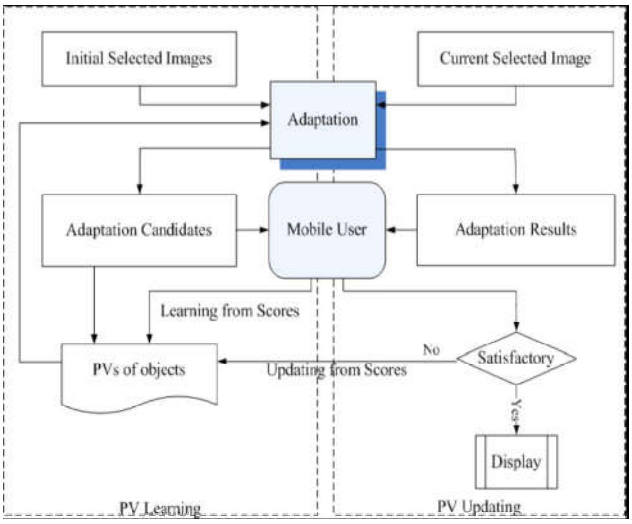


Fig 5. Feedback process for PV learning and updating

Feedback Process for PV Learning and Updating

Due to intrinsic characteristics of mobile device interfaces, simple and easy interaction schemes have to be designed for mobile users. For each grading, the user only needs to type a digit ranging from 1 to 9 to score whether it is consistent with his preference: 5 means no opinion; scores between 4 and 1 represent the degree of non-relevance, in which 4 is slightly non-relevant and 1 is highly non-relevant.

Degradation model for blurring image

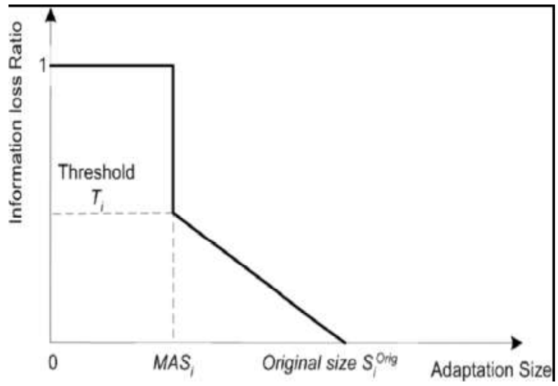
In degradation model for blurring image, the image is blurred using filters and an additive noise. The image can be degraded done by using Gaussian Filter and Gaussian Noise. Gaussian Filter represents the Point Spread Function which is a blurring function. The degraded image can be express by the equation $f = g * h + n$; Where * is the convolution operator, g is the clear image to recover, f is the observed blurred image, h is the blur kernel (or point spread function) and n is the noise.

Gaussian filter

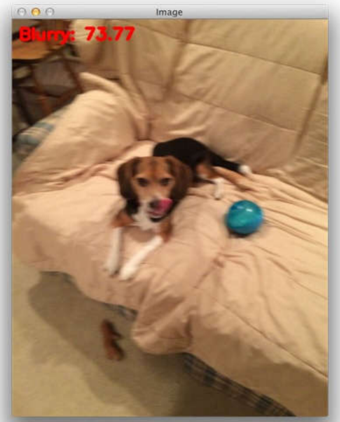
Gaussian filter is useful for blur an image by Gaussian function. It requires two specifications such as mean and variance. They are weighted blurring. Gaussian function is of the following form where σ is variance and x and y are the distance from the origin in the x axis and y axis respectively.

User Centric Semantic Adaptation

Given the key semantically meaningful objects contained in a consumer photo, the relevance of different objects for a given user preference can be varied. Moreover, the relevance of the same object for different mobile users may vary substantially Given these different PVs of different objects for various mobile users as well as the variety of mobile display capacities, the adaptation module has to decide what content to adopt adaptively according to these varying conditions. The goal of user centric adaptation is to simultaneously panelize the selection of contents not preferred by the user and preserve the user preferred objects with high quality depending on the degree of their relevance to user, under the limited mobile display constraints.



Experiments



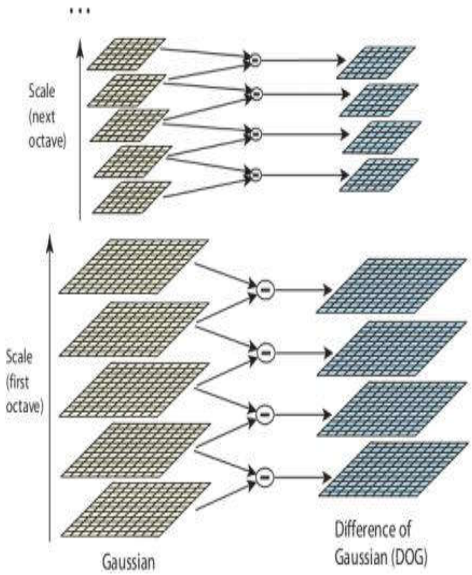
Detecting the amount of blur in an image using the variance of Laplacian. We implemented the variance of Laplacian method to give us a single floating point value to represent the "blurriness" of an image. This method is fast, simple, and easy to apply.

SIFT (Scale-Invariant Feature Transform)

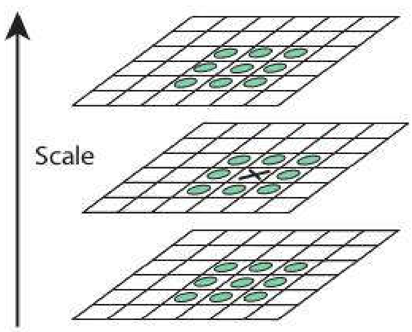
They are rotation-invariant, which means, even if the image is rotated, we can find the same corners. It is obvious because corners remain corners in rotated image also. But what about scaling? A corner may not be a corner if the image is scaled. For example, check a simple image below. A corner in a small image within a small window is flat when it is zoomed in the same window. So Harris corner is not scale invariant.

Scale-space Extreme Detection

From the image above, it is obvious that we can't use the same window to detect keypoints with different scale. It is OK with small corner. But to detect larger corners we need larger windows. For this, scale-space filtering is used. In it, Laplacian of Gaussian is found for the image with various σ values. This process is done for different octaves of the image in Gaussian Pyramid.



Once this Dog is found, images are searched for local extrema over scale and space. For e.g., one pixel in an image is compared with its 8 neighbors as well as 9 pixels in next scale and 9 pixels in previous scales. If it is a local extrema, it is a potential key point.



Key point Localization

Once potential key points locations are found, they have to be refined to get more accurate results. This threshold is called contrast Threshold in OpenCV. Dog has higher response for edges, so edges also need to be removed.

If this ratio is greater than a threshold, called edge Threshold in OpenCV, that key point is discarded. It is given as 10 in paper. So it eliminates any low-contrast key points and edge key points and what remain is strong interest points.

Key point Descriptor

Now key point descriptor is created. A 16x16 neighborhood around the key point is taken. It is divided into 16 sub-blocks of 4x4 sizes. For each sub-block, 8 bin orientation histogram is created. So a total of 128 bin values are available. It is represented as a vector to form key point descriptor. In addition to this, several measures are taken to achieve robustness against illumination changes, rotation etc.

Key point Matching

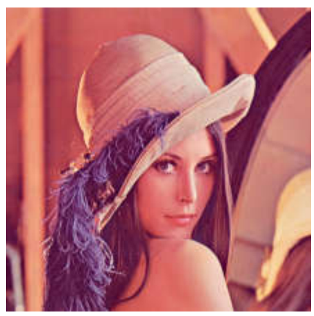
Key points between two images are matched by identifying their nearest neighbors. But in some cases, the second closest-match may be very near to the first. It may happen due to noise or some other reasons. In that case, ratio of closest-distance to second-closest distance is taken. If it is greater than 0.8, they are rejected. It eliminators around 90% of false matches while discards only 5% correct matches, as per the paper. So this is a summary of SIFT algorithm. For more details and understanding, reading the original paper is highly recommended. Remember one thing, this algorithm is patented.

SIFT in OpenCV

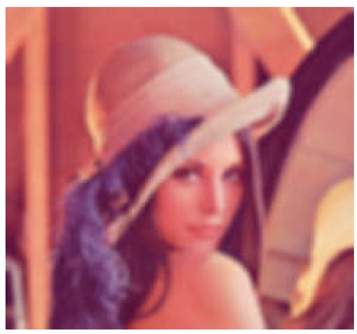
So now let's see SIFT functionalities available in OpenCV. Let's start with key point detection and draw them. First we have to construct a SIFT object. We can pass different parameters to it which is optional and they are well explained in docs. Sift. Detect () function finds the key point in the images.

RESULTS AND DISCUSSION

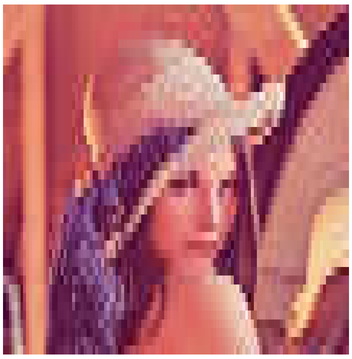
Original Image



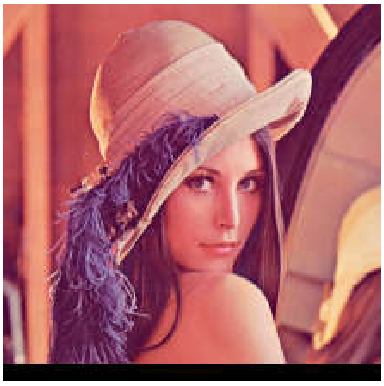
Blur image



Resized image



Shift image



Conclusion

In the above method first the image will be retrieved from the particular snap (or) album. That image will be correctly retrieved using manifold ranking of blocks. In that method the corresponding image will be split as number of pixels based on the color and texture. After that the image will be going under the Bayesian fusion algorithm for entire clearance of an image. This algorithm makes the image to be clear when we zoom that in mobile display devices. Normally while taking photos we couldn't get the image clearly which is shown back to the main image. That means if a person standing in front of the Tajmahal while taking a snap, at that time if we focus on the person, that person's image only will be captured clearly. That background (Tajmahal) image can't be captured clearly. That can be avoided here using recent and current techniques.

Acknowledgement

Open CV method first considers the original image. Then that image will be resized for detecting. Generally the image will be resized for detecting the blur part easily. Finally the blur part will be removed successfully. These all methods provides clear and unblurred image totally and even while zooming an image. This method having one drawback. We can get the clear image after zooming that. But while zooming, the image will not be clear. It will be looking like shaky image. After zooming only we can get the image with high resolution. So research is based on while we zooming also the image should be looking clear. While zooming we will move the particular of the image up and down, left and right.

REFERENCES

Chen, L. Q., Xie, X., Fan, X., Ma, W. Y., Zhang, H. J. and Q. Zhou, H. 2003. "A visual attention model for adapting image on small displays,", pp. 353–364, Oct.

Lagendijk, R. L. 2000. Basic Method for Image Restoration and identification, Academic Press.

Liu, H., Jiang, S., Huang, Q., Xu, C. and Gao, W. 2007. "Region-based visual attention analysis with its application in image browsing on small displays," pp. 305–308.

Mallat, S. and Hwang, W.L. 1992. "Singularity Detection and Processing with Wavelet," IEEE Trans. On Information Theory, March, pp.617-643.

Marichal, X., Ma, W.Y. and Zhang, H.J. "Blur Determination in the Compressed Domain Using DCT Information," Proceedings of the IEEE ICIP'99, pp. 386-390.

Suh, B., Ling, H., Bederson, B. B. and Jacobs, D. W. 2003. "Automatic thumbnail cropping and its effectiveness," in Proc. pp. 95–104.

****