Indonesian Journal of Electrical Engineering and Computer Science
Vol. 7, No. 1, July 2017, pp. 226 ~ 233
DOI: 10.11591/ijeecs.v7.i1.pp226-233

Content Based Image Retrieval using Feature Extraction in JPEG Domain and Genetic Algorithm

N. Parvin*[1], P. Kavitha[2]
[1]Department of computer science, Adhiyaman arts and Science College for women Srinivasa nagar, uthangarai, Tamilnadu, India
[2]Department of computer science, Loyala College of arts and science, Oilpatty, Rasipuram, Tamilnadu, India
*Corresponding author, e-mail: parvinfairose@gmail.com

Abstract

Content Based Image Retrieval (CBIR) aims at retrieving the images from the database based on the user query which is visual form rather than the traditional text form. The applications of CBIR extend from surveillance to remote sensing, medical imaging to weather forecasting, and security systems to historical research and so on. Though extensive research is made on content based image retrieval in the spatial domain, we have most images in the internet which is JPEG compressed which pushes the need for image retrieval in the compressed domain itself rather than decoding it to raw format before comparison and retrieval. This research addresses the need to retrieve the images from the database based on the features extracted from the compressed domain along with the application of genetic algorithm in improving the retrieval results. The research focuses on various features and their levels of impact on improving the precision and recall parameters of the CBIR system. Our experimentation results also indicate that the CBIR features in compressed domain along with the genetic algorithm usage improves the results considerably when compared with the literature techniques.

***Keywords**: DCT (Discrete Cosine Transform), GA (Genetic Algorithm), CH (Color Histogram), Color Moments; Precision and Recall*

1. Introduction

Image processing research is growing consistently and lot of new developments are there in the same [26-29]. Content based image retrieval is a database management system that is generally used for retrieval of images based on the image similarity with the user query image. It uses visual contents in order to search the images from a large database and is an active research area since 1990. There has been a remarkable progress in the past decade but most of the theoretical research and CBIR system development has happened only on the spatial domain. There are still many open problems to be solved that attract researchers from different disciplines. In a typical CBIR system as shown in Figure 1, the visual contents of the images are not directly compared; rather they are extracted and described in the form of multi-dimensional feature vectors. These vectors in turn form the feature database that is used in comparison and retrieval process. The user will provide the input in the form of an example image or as a sketch. The system will then extract the same set of features from this query image and compare it with all the features in the database to retrieve back the closest ones.

The same principle can be extended for both spatial domain as well compressed domain, the only difference being in the set of features extracted and the method of extraction.

In order to improve the results of the CBIR system, most works accommodates a feedback mechanism either explicit or implicit and use the end users response in predicting the outputs.

The visual contents of an image include color, texture and shape. These features can be extracted through different mechanisms that are discussed in depths across the literature [1]. But most of the research work focuses on extracting these features from the spatial domain on the RGB or HSV values directly [2].

Received March 29, 2017; Revised June 15, 2017; Accepted June 30, 2017

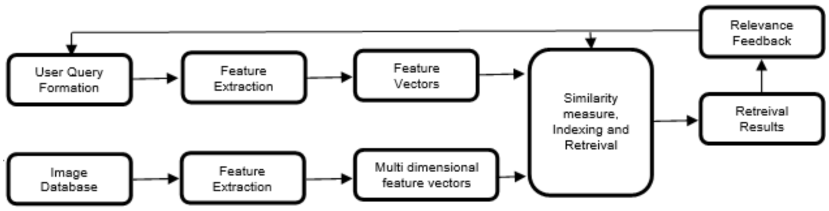

Figure 1. Content Based Image Retrieval System

According to Internet Trends 2014 code conference [3], people upload an average of 1.8 billion digital images every single day and that is equivalent to 657 billion photos per year which is a very huge number. Most of these images are in JPEG compressed format and hence a direct retreival mechanism in this compressed domain would be better as compared to decompressing them completely which is a costly process in digital world. In this work, we have identified the features that can be extracted in the compressed domain on the DCT coefficients itself and also propose to use genetic algorithms for getting the relevance feedback from the end user and apply it in fine tuning the retrieved results.

2. Literature Survey

Image content includes both visual and semantic content for its description. It could also be specific to a particular domain or a generic feature. Specific ones include human face, finger prints etc. are used in specific application while the generic ones like color and texture are used in other domains. The challenge involved in designing a robust CBIR system is in extracting the content descriptors which should be invariant to the issues introduced due to imaging process.

Researchers from different fields including computer vision, database management, and information retreival are all attracted to CBIR problem statement [5] from the year 1992. Also since the year 1997, the different research publications are mostly based on primitive features extracted from the spatial domain [6] and some had relevance feedback added to it [7].

Efforts were also made to combine the visual features together for an improved and robust performance of CBIR system [8] but lacked the optimization solution as it takes a huge amount of time for feature extraction, comparison and retreival process [9]. This made the researchers to turn towards extracting the features in the compressed domain itself as the huge amount of data in the internet as well as local storage is in JPEG compressed format [12].

Color is an important feature in judging the image similarity and hence multiple research papers are published using this feature [17]. Texture is the next commonly used feature [19] as it helps in identifying the repeated patterns in the image or a video frame. To differentiate between different objects of the same color, Texture will be very useful. The different texture descriptors that are usually extracted from the images include mean, standard deviation, angular second moment, inverse difference moment, sum average, contrast, correlations and sum variance. These above mentioned descriptors have been used to select on the number of uniform or textured or coarse blocks present in the image.

Shape is the last used feature in most CBIR systems. In order to find the shape of the image, different edge detection algorithms [14] are used. They compute the moment of the query image and integrate the moment features of given images row and column wise [21]. Then they are compared against the database stored images moments and the closest matching ones are retrieved back.

For two reasons including the overhead in decompressing the images and the fact that the compressed ones hold the most wanted information by removing the redundant information during the compression process, there were research works on compressed domain CBIR systems as well [22]. Zhe-Ming Lu and others [23] feel that the existing CBIR methods are far from the practical application and have presented a novel approach in extracting the color, texture and other features on the DCT domain. Their system is found to be suitable for all color images of different sizes as well.

Genetic algorithms (GA) exist in the literature and use as well for a while and is a method for explaining both constrained and unconstrained optimization problems based on a natural selection process. The algorithm modifies a population repeatedly of individual solutions. Genetic algorithms along with precision and recall parameters can be used as a fitness function in selecting the optimum weights of features [24]. It is observed that the optimal weights of features that are computed by GA have improved meaningfully all the assessment measures including average precision and recall for the combined features method. Though it is explored mostly in spatial domain features, we have extended the GA to the features extracted in the compressed domain and found the results to be even better.

3. Proposed Research Framework

Image Retreival can be classified based on different levels of complexity in extracting the features as Level 1, Level 2 and Level 3 respectively. Level 1 deals with primitive features including color, texture and shape while level 2 is based on logical or derived features and level 3 is about retrieval based on abstract attributes. Level 3 takes in to account the semantics, purpose and content of the image on top of primitive features. Semantic gap deals with the difference between the user query and the retrieval results. Tougher the feature extraction from a meaningful data will reduce the semantic gap. For this purpose, we will first extract the features from the DCT domain which has no redundant information in it.

3.1. JPEG Standard and Content Analysis

JPEG (Joint Photographic Experts Group) is a lossy compression technique used in storage of digital images. The degree of compression is an adjustable parameter based on the application and use cases. Moreover it is the most common format used for storing and transmitting the images on the World Wide Web [25]. So, we consider JPEG images compared to other compression formats in this research work.

JPEG uses a lossy compression technique based on the Discrete Cosine Transform (DCT). This transform converts a 2D image in the spatial domain to the frequency domain. As the human psychovisual system does not give importance to the high frequency information, this transformation in the process of quantization removes all these redundant information and stores back only the required details to reproduce the image. Though this method comes under lossy compression technique, we get the required amount information that is needed for the content based image retreival system.

The encoding process involves color space transformation, down sampling, block splitting followed by discrete cosine transformation of the data, quantization, entropy encoding and packing. The decoder follows the exact reverse process as shown in Figure 2 and in this work, we concentrate hugely on DCT part as the feature extraction is surrounding that. The two dimensional DCT is given by:

$$G_{u,v} = \frac{1}{4}\alpha(u)\alpha(v)\sum_{x=0}^{7}\sum_{y=0}^{7} g_{x,y}\cos\left[\frac{(2x+1)u\pi}{16}\right]\cos\left[\frac{(2y+1)v\pi}{16}\right]$$

In the above equation 'u' represents the horizontal spatial frequency, v represents the vertical spatial frequency. g(x,y) corresponds to pixel value at co-ordinates (x,y) and G(u,v) represents the DCT coefficients at co-ordinates (u,v).

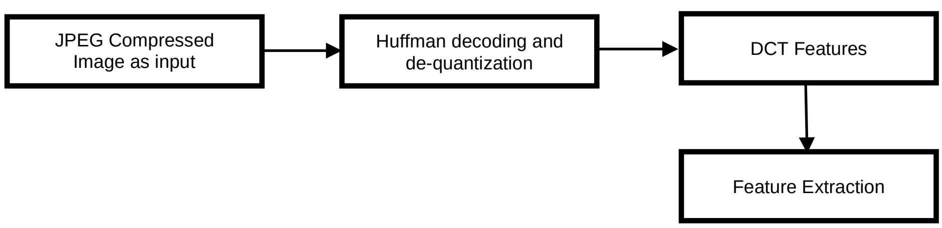

Figure 2. Feature Extraction from JPEG compressed image

At the end of DCT transform, we get 64 frequency components. The left most and the first DCT coefficient represents the DC component which corresponds to the average luminance in this block and is the most important information. The other 63 coefficients can be classified as the low and high frequency components based on their position in the block. Various features can be extracted from these coefficients which are discussed in the next section.

4. Feature Extraction in the DCT Domain

We extract the features from the JPEG compressed image before proceeding with image matching and retreival process. In this section we detail the different features considered in our work for image retrieval. The different features include color and texture information from the compressed images.

4.1. Color Feature from DCT

Color is important visual information and it was established using the content based image indexing. So we get the color moments as the first feature vector in our work as well. It is represented through the first two moments namely the mean and the standard deviation.

$$E = \frac{1}{N \times M}\sum_{x=1}^{N}\sum_{y=1}^{M} p(x,y),\ \sigma = \left(\frac{1}{N \times M}\sum_{x=1}^{N}\sum_{y=1}^{M}\left(p(x,y)-E\right)^2\right)^{\frac{1}{2}}$$

Along with the color moments, we also use color histogram for identifying the image content. This is given by:

$h(i) = n_i / n$

Here n_i represents the number of pixels of the ith color and n represents the number of pixels in the image. We extract and store these feature vectors in the database. We calculate them independently for every block identified and store all of them.

4.2. Texture Feature from DCT

Texture is the next feature identified for image retreival purposes. The first DC coefficient in the DCT block represents the energy information of the image while the remaining AC components indicate the frequency information. The direction information is also present in the block. The dominant directions along with the gray level variations of the image are extracted and stored in the image database as independent feature vectors. Each sub block is processed independently as shown in Figure 3.

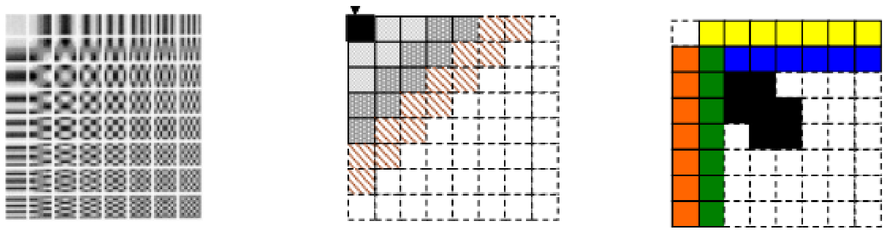

Figure 3. Texture Extraction from DCT blocks

The vertical and horizontal information's are extracted from each block and processed for further usage. Once all the values are computed, we calculate the mean and standard deviation over these values and store them. The feature vectors represent the average greyness, horizontal texture, vertical texture and diagonal texture of a particular image sub block, respectively. These coefficients are intentionally selected as they are vital to the greyness and image directionality. Also, only the low frequency coefficients of the sub block are selected as they deliver higher energy level in a usual DCT domain.

5. Genetic Algorithm in CBIR Compressed Domain Approach

The difference seen between the user's need and the images retrieved represents the semantic gap in CBIR systems. In order to reduce this semantic gap, relevance feedback mechanisms are introduced. By means of this relevance feedback, we will be able to integrate human perception in to the query and evaluate the results as well.

Genetic algorithms which is a method for solving optimization problems can thus be borrowed from biological evolution to relevance feedback mechanism in CBIR systems. Genetic algorithms simulates the process of evolution and evolution is an optimizing process. Through this process, the successive generations will become better and better and each generation is like an iteration in numerical methods. In CBIR system also, we improve the results over every iteration through relevance feedback from the end user.

Selection, crossover operation and mutation along with acceptance are the different stages involved in genetic algorithm as depicted in Figure 4. Selection operation is the first in the list which takes care of selecting the individual feature vectors as parents that can generate offspring. The selection happens through the fitness level. The better the feature vectors are, there are more chances for them to be selected. The successor's generation in a genetic algorithm is found by a set of operators that recombine and mutate certain members of the existing population. Crossover and mutation are the two most common operators used for this. The crossover operator yields two new offspring from two parent strings, by replicating certain bits from each parent. The mutation on the other hand will make small changes to the bit string by selecting a random bit and modifying its value.

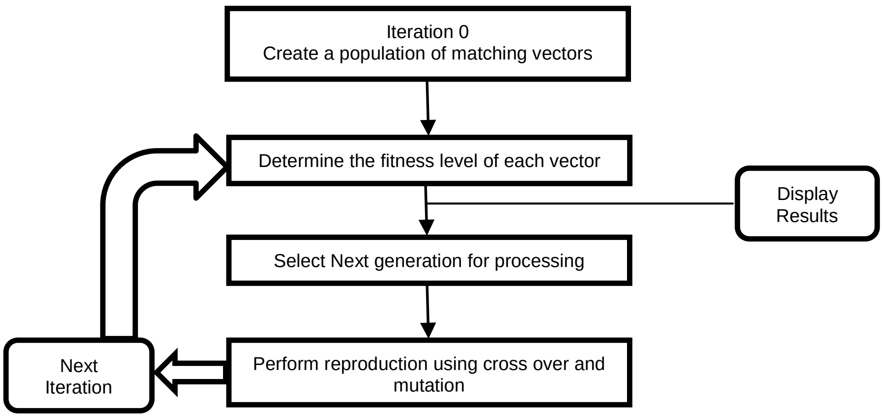

Figure 4. Genetic Algorithm in refining the CBIR results

The fitness function is obtained by the following equation:

$$F_1 = R_r - R_n - N_r$$

where R_r is the number of images retrieved that are in the set of relevant images provided by the database and judged as relevant by the user implicitly without their knowledge. R_n refers to the number of images retrieved that are not included in the standard relevant set and N_r is the number of relevant documents that is not retrieved.

The images are sorted now in the order of decreasing similarity and the final fitness is given by the equation:

$$F = \frac{1}{|D|}\sum_{i=1}^{|D|}\left(r(d_i)\sum_{j=i}^{|D|}\frac{1}{j}\right)$$

where |D| refers to the total number of images retrieved and r(d) signifies the relevance of the images d. This will be unity if the image is relevant otherwise zero. By this way, during every iteration we will be able to filter out the irrelevant images and display back the matching ones.

6. Results and Discussion

One of the main goals of this research was to improve the retreival accuracy of the CBIR system and find its use in multiple fields. We have used the STL-10 dataset for our experimentations. It is motivated by the CIFAR-10 dataset but has alterations to it. We have 10 different classes including the airplane, tiger, bird, duck, players, buildings, flowers, ships and water images in it. We have used 500 images in the database and 25 images for testing. All the images are of 96*96 pixels and JPEG compressed.

As a first step, we have extracted the color and texture features from the 500 images that are present in our database and stored them as multidimensional feature vectors. Then we had a MATLAB implementation of GUI for querying a test image and to display the results. We started testing the 25 images one by one. Every time a query image is provided, we extract the same set of features from it and used Euclidean distance method to find the distance between the extracted feature vector value and the data present in the database. We then used precision and recall parameters for estimating our retrieval accuracy.

As a final step, we introduced the Genetic algorithms in getting the relevance feedback from the end user and used it in fine tuning the retreival results. It is clear from the results that the retrieval accuracy increases by 12% when we use relevance feedback and genetic algorithms in compressed domain content based image retreival systems. The data set along with the retrieved images are illustrated in Figure 5 while the results are tabulated in Table 1.

(a) (b) (c)

Figure 5. Snapshot of the data set (a), input test image (b) and retrieved results (c)

Though the original retreival results from the above Figure 5 shows only the first four to be relevant, we find that with the help of relevance feedback and genetic algorithms, the retreival accuracy improves considerably and those results are tabulated in Table 1. Precision found in the table is defined as number of retrieved images that are relevant divided by total number of retrieved images from the database. The other parameter recall is defined in the similar way as the number of retrieved relevant images from the database divided by total number of relevant images.

Table 1. Recall and Precision Ratio for Different Kinds of Images

Images	Recall Ratio (%)	Precision Ratio (%)
Flower	85	92
Tiger	72	85
Duck	60	78
Building	74	80
Players	86	90
Bird	78	88
Water	82	86
Elephant	76	92
Sun	80	88

We have presented two stage retrieval processes in this research work. At the first stage, we have used the primitive features namely the color and texture which are extracted from the DCT coefficients of the JPEG compressed image and at the second stage we have used the genetic algorithm in fine tuning the retrieved results. We find that the results improve a lot with relevance feedback from the end user when genetic algorithms are deployed in compressed domain content based image retrieval as compared to the spatial domain content based image retrieval systems.

7. Future Recommendations

The proposed idea can be extended towards DWT compressed images as used in JPEG 2000 methodology. Though we have tested against an animal data set, the proposed approach can be tested against different commercial applications including crime prevention, security checks, medical diagnosis, intellectual properties, architectural and engineering designs etc. to lost a few. The field of open questions is still very huge and much interesting work can be done to get a large and robust feature set in content based image retrieval systems.

References

[1] T. Dharani and I. Laurence Aroquiaraj. *A survey on content based image retrieval.* International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME), 2013.

[2] Jun Yue, Zhenbo Li, Lu Liu and Zetian Fu. *Content-based image retrieval using color and texture fused features.* Mathematical and Computer Modelling. Vol 54: 3–4, August 2011, Pages 1121–1127.

[3] Mary Meeker. *Internet Trends 2014 – Code Conference.* http://www.kpcb.com/internet-trends, June 1, 2016.

[4] Ricardo Da Silva Torres and Alexandre Xavier Falcão. *Content-Based Image Retrieval: Theory and Applications*, RITA, Volume XIII, Número 2, 2006.

[5] Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta and Ramesh Jain, *Content-Based Image Retrieval at the End of the Early Years*, IEEE Transactions on Pattern Analysis and Machine Intelligence archive, Volume 22 Issue 12, December 2000, Page 1349-1380

[6] N. Vasoncelos, and A. Lippman, *A probabilistic architecture for content-based image retrieval*, Proc. Computer vision and pattern recognition, pp. 216-221, 2000.

[7] Su Z, Zhang H, Li S and Ma S. *Relevance feedback in content-based image retrieval: Bayesian framework, feature subspaces, and progressive learning*, IEEE Trans Image Process. 2003; 12(8):924-37.

[8] Rashid Ali, S.M. Zakariya and Nesar Ahmad, *Combining Visual Features of an Image at Different Precision Value of Unsupervised Content Based Image Retrieval*, IEEE 2010.

[9] K. Satya Sai Prakash and RMD. Sundaram, *Combining Novel features for Content Based Image Retrieval*, 14th International Workshop on Systems, Signals and Image Processing, 2007 and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services.

[10] Ebroul Izuierdo and Qianni Zhang, Senior Member, IEEE, *Histology Image Retrieval in Optimized Multifeature Spaces*, IEEE Journal of Biomedical and Health Informatics, Vol. 17, No 1, January 2013

[11] J. Assfalg, A. D. Bimbo, and P. Pala, *Using multiple examples for content-based retrieval*, Proc. Int'l Conf. Multimedia and Expo, 2000.

[12] Padmashri Suresh, R. M. D. Sundaram and Aravindhan Arumugam, *Feature Extraction in Compressed Domain for Content Based Image Retrieval*, International Conference on Advanced Computer Theory and Engineering, 2008. ICACTE '08.

[13] Gerald Schaefer, *Pixel domain and compressed domain image retrieval features*, Eighth International Conference on Digital Information Management (ICDIM), 2013.

[14] Abhishek, S.N., Shriram K Vasudevan and Sundaram, R.M.D., *An enhanced and efficient algorithm for faster, better and accurate edge detection*, Communications in Computer and Information Science, Springer Verlag, Volume 679, p.24-41 (2016).

[15] M. Antonelli, S. G. Dellepiane, and M. Goccia, *Design and implementation of Web-based systems for image segmentation and CBIR*, IEEETrans. Instrum. Meas., vol. 55, no. 6, pp. 1869–1877, Dec. 2006.

[16] W. H. Hsu, L. S. Kennedy and S.-F. Chang. *Reranking Methods for Visual Search*. IEEE Multimedia, vol. 14, no. 3, pp. 14-22, 2007

[17] Remo c. Veltcamp, Mirela Tanase, and Danielle Sent. *Features in content based image retrieval systems*: A survey. In Remo C. Veltkamp, Hans Burkhardt, and Hans-Peter Kriegel, editors, State-of-the-Art in Content Based Image and Video Retrieval, chapter 5, pages 97–124. Kluwer Academic Publishers, Utrecht University, Department of Computer Science, Utrecht, the Netherlands, 2002.

[18] B. Chellaprabha and RMD. Sundaram, *Effective Video Streaming Using Reset-Controlled DCCP along with Data Quality Analysis*, Asian Journal of Information Technology, Year: 2016 | Volume: 15 | Issue: 20 | Page No.: 3986-3990
[19] Shankar M. Patil, *Content Based Image Retreival using Color, Texture and Shape*, International Journal of Computer Science & Engineering Technology (IJCSET), Vol. 3 No. 9 Sep 2012.
[20] G. Beligiannis, L. Skarlas, and S. Likothanassis, *A generic applied evolutionaryhybrid technique for adaptive system modeling and information mining*, IEEE Signal Process. Mag.—Special Issue on —Signal Processing for Mining Information‖, vol. 21, no. 3, pp. 28–38, May 2004.
[21] Reshma Chaudhari and A. M. Patil, *Content Based Image Retrieval Using Color and Shape Features,* International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 1, Issue 5, November 2012.
[22] Xia Wu, *Based on the Characteristics of the JPEG Image Retrieval System Research*, Seventh International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), 2015.
[23] Zhe-Ming Lu, Su-Zhi Li and Hans Burkhardt, *A Content based image retreival scheme in JPEG compressed domain,* International Journal of Innovative Computing, Information and Control 1349-4198 Volume 2, Number 4, August 2006.
[24] Raghupathi Gali, M. L. Dewal and R. S. Anand, *Genetic Algorithm for Content Based Image Retrieval*, Fourth International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN), 2012.
[25] Bin Li, Tian-Tsong Ng, Xiaolong Li, Shunquan Tan and Jiwu Huang, *Revealing the Trace of High-Quality JPEG Compression Through Quantization Noise Analysis*, IEEE Transactions on Information Forensics and Security, Volume: 10, Issue: 3, March 2015.
[26] Shriram, K.V., Priyadarsini, P.L.K. and Baskar, A., 2015. An intelligent system of content-based image retrieval for crime investigation. *International Journal of Advanced Intelligence Paradigms*, *7*(3-4), pp.264-279.
[27] Vasudevan, Shriram K., Karthik Venkatachalam, S. Anandaram, and Anish J. Menon. A Novel Method for Circuit Recognition through Image Processing Techniques. *Asian Journal of Information Technology* 15, no. 7 (2016): 1146-1150.
[28] Shriram, K. V., P. L. K. Priyadarsini, M. Gokul, V. Subashri, and R. Sivaraman. A novel approach using THES methodology for CBIR. In *Signal Processing Image Processing & Pattern Recognition (ICSIPR), 2013 International Conference on*, pp. 10-13. IEEE, 2013.
[29] Parthasarathi, V., M. Surya, B. Akshay, K. Murali Siva, and Shriram K. Vasudevan. Smart control of traffic signal system using image processing. *Indian Journal of Science and Technology* 8, no. 16 (2015): 1.