

Urban Feature Classification Technique from RGB Data using Sequential Methods

Hassan Elhifnawy
Civil Engineering Department
Military Technical College
Cairo, Egypt

Abstract- This research produces a fully feature classification technique from RGB image. The proposed classification technique makes a combination from different segmentation methods to classify all features in urban areas. RGB color channels are used to produce two color invariant images. The first one is produced using blue and green color channels for vegetation areas identification. The color invariant image is segmented into two clusters; one of them represents vegetation areas. The second color invariant image is produced using all RGB color channels. Otsu segmentation technique is applied to detect shadows areas as one of the resulting clusters. However, color space is not suitable for detecting roads and buildings. Consequently, it is transformed into two other color spaces. Luminance color channel is extracted from first color space, and hue and saturation color channels are extracted from the other one. This is followed by global color thresholding technique on these color channels individually and together to detect roads, sandy or unhealthy vegetation areas. The remaining features of the original RGB image of urban area of study are classified as buildings.

The investigated technique is automated and suitable for all urban areas from high resolution RGB images captured by digital cameras or satellite imaging sensors. This technique is performed exceptionally when there are high distinguishable texture properties for different features. It will be shown that, in case of existing texture similarity properties of different features, it is recommended to use additional information such as spatial data to be able to classify them efficiently.

Keywords—RGB, color invariant, segmentation, urban features, classification.

I. INTRODUCTION

The objective of this research, as shown in Fig. 1, is to get a suitable method used for evaluating effects of environmental hazards, as earthquakes and floods, on urban area from aerial RGB images. This objective is a motivation for searching about feature extraction technique for extracting urban area features from commercial RGB images. After extracting urban area features, it is possible to determine the areas of affected features and evaluate the cost of repairing or reconstruction.

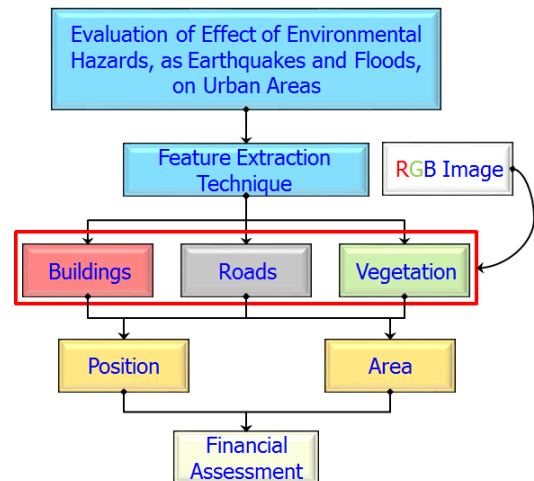


Fig. 1 Research Objective

RGB images provide rich semantic information for ground and non-ground objects [1]. Several researches are performed to extract features from airborne RGB or satellite images. One of these research papers introduced a fully automated technique for road extraction from satellite image [2]. This technique is based on a combination of A trous Wavelet Transform (AWT) algorithm, fuzzy logic, and Hough transform to detect road candidates. Wavelet filtering based on A trous Wavelet Transform (AWT) algorithm is applied on images in gray scale format using two wavelet base functions, Haar and db8, and stopped at fourth level of decomposition. The resulting two images are fused in one image using Karhounen-Louve Transform (KLT). Fuzzy logic and Hough transform are used for building fuzzy logic interface algorithm to detect road candidates. Road identification is applied for each window after dividing fused image into small windows. This combined technique provides acceptable results in case of non-urban areas, but for urban areas some of the buildings are extracted as road candidates. Sirmacek and Unsalan [3] detected building roofs using different approach based on calculating color invariant image from red and green color channels. This technique is successful in case of red roof buildings only and it is not suitable for any other roof color. Shorter and Kasparis [4] investigated an automatic technique for building extraction from RGB images captured by digital cameras. The image is segmented by using color segmentation process by removing small areas based on specific area thresholding value. The segmented image is used for producing color invariant images based on green and blue color channels for vegetation areas identification. The raw RGB images are used for producing color invariant images for shadows area identification based

on all color channels. A Watershed segmentation technique was applied on raw RGB images followed by calculating solidity properties for all segmented regions to investigate building and non-building candidates. The building image is investigated after removing vegetation and shadows areas from building candidate image. This technique succeeded in detecting vegetation, shadows and building areas from RGB image, but it depends on the input image and color varieties among these different features. Although, RGB images provide rich semantic information for ground and non-ground objects, RGB color space is not suitable for extracting all features. Song and Shan [5] investigated a building extraction technique from high resolution RGB images. The RGB images are transformed to CIE L* a* b* color space. Active contour image segmentation is applied to detect building boundaries, and JSEG frame worm is used to construct building polygons and 3-D wired frame. This technique succeeded in case of red rooftop buildings that provide high contrast between buildings and background. Bong et al. [6] investigated different color channels ranges representing roads texture property after transforming RGB image from *RGB* color space to *YCbCr* and *HSV* color spaces. Road images are produced by applying global thresholding on different color channels by specific ranges. This technique is reliable for high resolution satellite images, but with clearly differentiable semantic information between buildings and roads to avoid extracting building pixels as road candidates. This technique gave acceptable results for roads detection, but there are no color channels ranges suitable for detecting all different features. As a conclusion for the previous work done, that an efficient combination between supervised and unsupervised techniques will be important in having a complete automated feature classification technique for urban areas. Vegetation and shadows detection is based on segmenting color invariant images based on natural properties of the gray level values and is called unsupervised segmentation technique. Roads detection is based on global color thresholding for different color channels using specific values and is called supervised segmentation technique. This research paper produces a fully features classification technique from RGB image by combining the two classification technique efficiently, which helps in the efficient automated classification of the above mentioned features (vegetation, roads, shadows and buildings) from RGB images.

II. FEATURE CLASSIFICATION ALGORITHM

The proposed classification technique is composed of two main stages as shown in Fig. 2. The first stage is the color segmentation using unsupervised segmentation method to identify shadows and vegetation areas. RGB color channels are used to produce two color invariant images for vegetation and shadows areas identification. Otsu, unsupervised segmentation technique, is applied on these color invariant images to extract shadows and vegetation candidates. The second stage is the color segmentation using supervised segmentation technique to identify roads and buildings. Input RGB image (from the first step) is transformed into two different color spaces, *YCbCr* and *HSV*. Luminance color channel is extracted from *YCbCr* color space, and hue and saturation color channels are extracted from *HSV* color space. Global color

thresholding is applied on all color channels to identify road candidates.

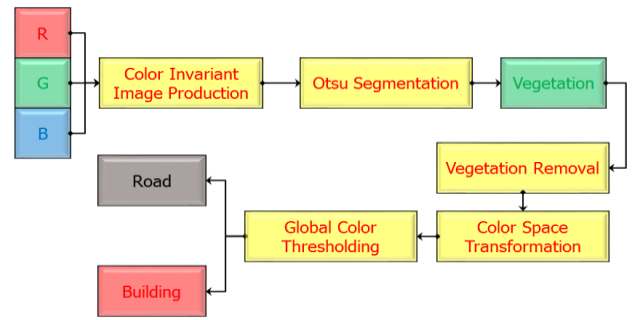


Fig. 2 Feature Extraction Algorithm

III. METHODOLOGIES

This section introduces an overview about segmentation methodology used for detecting vegetation and shadow that called Otsu segmentation “References [7, 8]” from two different color invariant images. Also, it introduces a mathematical model for color transformation from RGB color space to two different color spaces used for detecting road candidates.

A. Otsu Segmentation

Otsu thresholding is an unsupervised segmentation method. It segments an input image into two main clusters. This method is based on having a statistical histogram of the image gray values. The selected gray thresholding value k has to maximize between class variance represented in (1) [7, 8].

$$\sigma_B^2 = \omega_0(u_0 - u_T)^2 + \omega_1(u_1 - u_T)^2 \quad (1)$$

where:

$$\omega_0 = \sum_{r=1}^k P_r$$

$$\omega_1 = \sum_{r=k+1}^L P_r$$

$$u_0 = \sum_{r=1}^k rP_r / \omega_0$$

$$u_1 = \sum_{r=k+1}^L rP_r / \omega_1$$

$$u_T = \sum_{r=1}^L rP_r$$

$P_r \geq 0$ = probability of existence of gray level r

B. Color Transformation

Equation (2) shows the mathematical model for color space transformation from *RGB* color space to *YCbCr* color space where Y refers to luminance color channel representing gray scale information, Cb and Cr components represent color different between blue and red channels and the reference value respectively. Equation (3) shows the mathematical model for color transformation from *RGB* color space to *HSV* color space [7, 9]. In *HSV*, H represents hue color channel that represents a true color, S represents saturation color channel as a degradation

measurements for diluting a true color by a white light, V is a value color channel, but it is not suitable for human interpretation, so the intensity value I is used instead of V color channel.

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.699 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2)$$

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 & \text{if } B > G \end{cases}$$

$$\theta = \cos^{-1} \left[\frac{\frac{1}{2} [(R-G) + (R-B)]}{\sqrt{[(R-G)^2 + (R-B)(G-B)]}} \right] \quad (3)$$

$$S = 1 - \frac{3}{(R+G+B)} [\min(R,G,B)]$$

$$I = \frac{1}{3} (R+G+B)$$

IV. URBAN FEATURE DETECTION

A. Shadows and Vegetation Detection

Fig. 3 shows the used RGB image in verifying the proposed technique.



Fig. 3 Input Image [Image Courtesy of Twisted Sifter]

Shadows and vegetation areas can be detected through color invariant images that represent ratios among different color channels in RGB color space directly [4]. Shadows can be detected from color invariant images using all RGB color channels shown in (4) [3], where SI represents shadows area candidates, R , G , and B represent RGB color channels. Vegetation areas can be detected from color invariant images using green and blue color channels from RGB color space shown in (5) [10], where VI represents vegetation candidates. Fig. 4 shows color invariant images and shadows and vegetation candidates after applying Otsu segmentation technique.

$$SI = \frac{4}{\pi} \arctan \left[\frac{R - \sqrt{R^2 + G^2 + B^2}}{R + \sqrt{R^2 + G^2 + B^2}} \right] \quad (4)$$

$$VI = \frac{4}{\pi} \arctan \left(\frac{G-B}{G+B} \right) \quad (5)$$

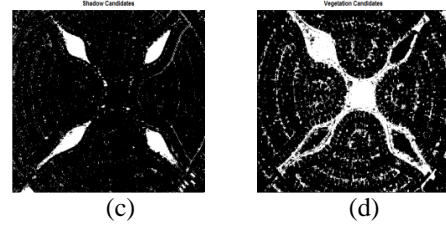
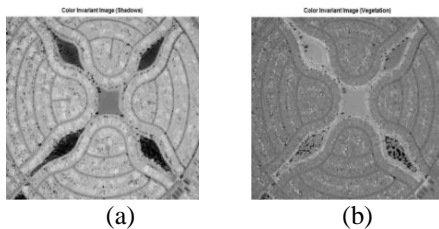


Fig. 4 Shadows and Vegetation Identification

B. Road Detection

Fig. 5 shows the input RGB image after removing vegetation and shadows. This image is considered as an input image for the second processing step.

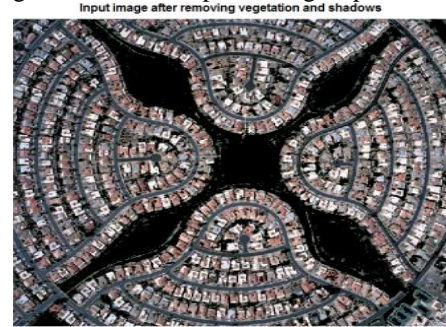


Fig. 5 Image after Vegetation and Shadows Removal

The used RGB image is transformed into two different color spaces ($YCbCr$ and HSV). Fig. 6 (a) shows the transformed RGB image in $YCbCr$ color space and Fig. 6 (b) shows the same image in HSV color space. The most important color channels used in extracting roads and sandy areas are luminance, saturation and hue. Fig. 7 shows the color channels used in global color thresholding process.

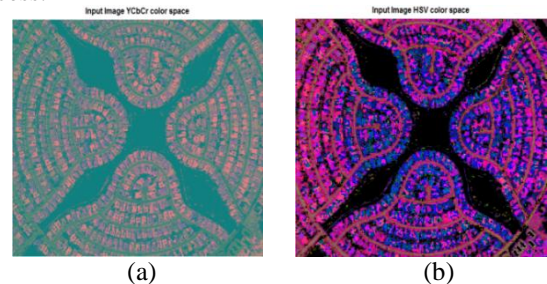


Fig. 6 Input Image in Two Different Color Spaces



Fig. 7 Color Channels for Road and Sandy Areas

Luminance color channel is used for extracting most probable road candidates. Equation (6) represents the mathematical model for road candidates identification, where RIL represents the road candidates and Y the luminance color channel. Fig. 8 show the extracted road image using luminance color channel.

$$RIL = \begin{cases} 1 & \text{if } Y \leq 100 \\ 0 & \text{Otherwise} \end{cases} \quad (6)$$

There are many areas other than roads still exist in the image and extracted as a road candidates (false).

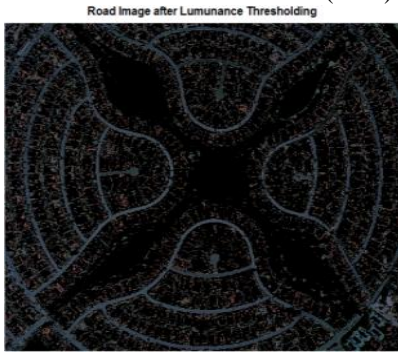


Fig. 8 Road Detection from Luminance Color Channel

Global color thresholding is applied on hue and saturation color channels to detect pixels defined as non road candidates, as shown in (7) and (8), where *NRH* and *NRC* represents the non-road candidates.

$$NRH = \begin{cases} 1 & \text{if } H < 0.05 \\ 0 & \text{Otherwise} \end{cases} \quad (7)$$

Fig. 9 shows the non-road RGB image after thresholding hue and all color channels.

$$NRC = \begin{cases} 1 & \text{if } L \leq 100 \text{ and } S \geq 0.03 \text{ and } H < 0.05 \\ 0 & \text{Otherwise} \end{cases} \quad (8)$$

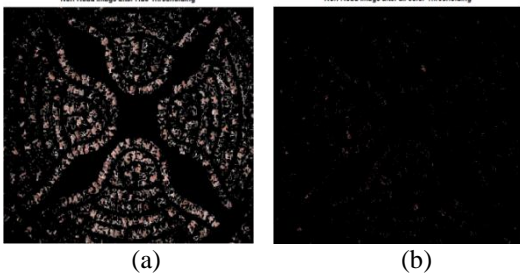


Fig. 9 Non- Road Images from Hue and all Color Channels

After studying several areas, urban and rural, it is found that sandy soil areas can be detected by applying global thresholding for all color channels using the proposed mathematical model shown in (9) where *SA* represents candidates of sandy areas.

$$SA = \begin{cases} 1 & \text{if } 110 \leq L \leq 160 \text{ AND } 0.10 \leq S \leq 0.25 \text{ AND } 0.05 \leq H \leq 0.20 \\ 0 & \text{Otherwise} \end{cases} \quad (10)$$

Fig. 10 (a) shows sandy soil candidates and Fig. 10 (b) shows RGB image of sandy areas. This mathematical model not only used for sandy areas identification, but also for detecting non-healthy vegetation areas.

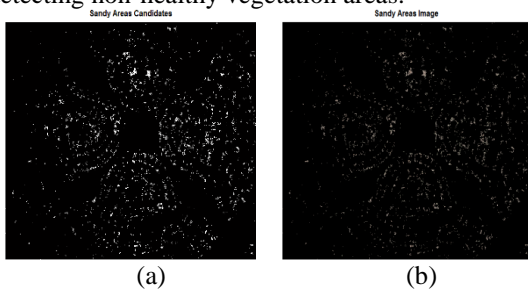


Fig. 10 Sandy Areas Identification

Now it is possible to extract road candidates from all previous thresholding by taking *RIL* as main road

candidates in RGB image and eliminating any other non-road candidates that extracted from *NRH*, *NRC*, and *SA*. The investigated mathematical model is shown in (11) where *RC* represents the road candidates from RGB image. Fig. 11 shows the extracted road candidates and the extracted RGB road image.

$$RC = \begin{cases} 0 & \text{if } NRH \neq 0 \text{ or } NRC \neq 0 \text{ or } SA \neq 0 \\ RIL & \text{Otherwise} \end{cases} \quad (11)$$

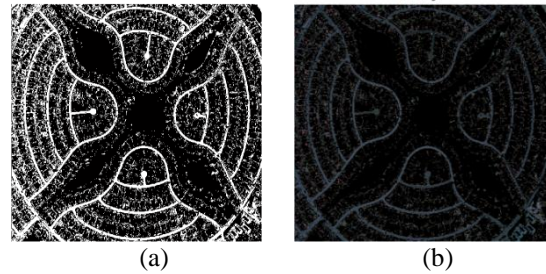


Fig. 11 Road Identification

C. Building Detection

The remaining features in the RGB image are classified as buildings. Building candidates are the used RGB image after eliminating all vegetation, shadows, sandy soil, and road candidates as using (12), where *BC* represents building candidates. Fig. 12 shows building candidates and the extracted RGB building image.

$$BC = \begin{cases} 0 & \text{if } VI \neq 0 \text{ or } SI \neq 0 \text{ or } SA \neq 0 \text{ or } RC \neq 0 \\ 1 & \text{Otherwise} \end{cases} \quad (12)$$

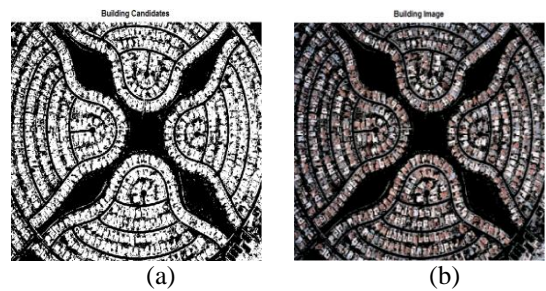


Fig. 12 Building Identification

V. URBAN AREA CLACSSIFICATION

Fig. 13 shows the final feature classification results for this urban area. Red color represents buildings, green color represents vegetation areas, black color represents roads, yellow color represents sandy areas, and gray color represents shadows.

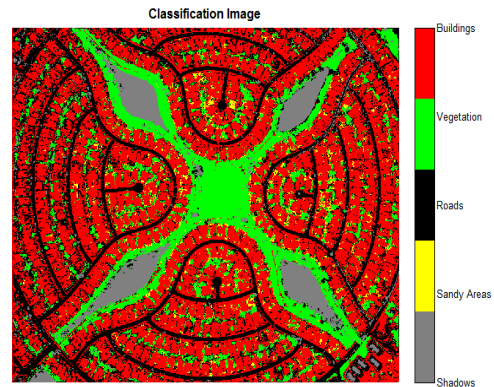


Fig. 13 Urban Area Classification

VI. ASSESSMENT OF CLASIFICATION TECHNIQUE

The objective of this research is to get an optimum result by using commercial RGB aerial images. The input image that is used in this research is from an open source through internet. So, there is no database for the area of study to assess the results of the proposed feature extraction technique. A supervised classification is applied on the input RGB images of area of study. The results of supervised classification are using to form a database for area of study. Fig. 14 shows the classification results after applying supervised classification.

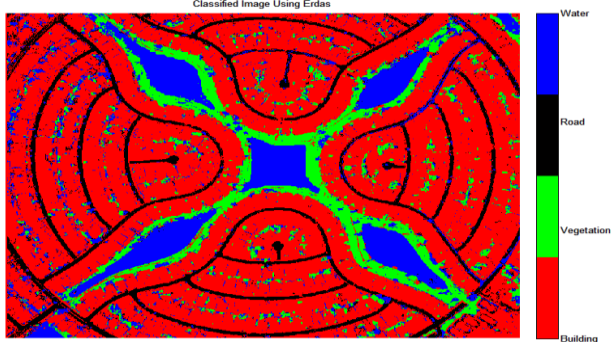


Fig. 14 Supervised classification results

Supervised classification is applied to classify the urban features buildings, vegetation areas and roads. The water areas are detected in the image of area of interest. So, the water area is classified in addition to the three urban features. The classification image, Fig. 14, is considered as a reference for classification results for the area of study. The comparative study is applied between results from proposed feature extraction technique and the results of classification after applying supervised classification on image of area of interest. The classification results are listed in form of confusion matrix [11] as shown in Table 1. The matrix represents the number of pixels for extracted features from proposed feature extraction technique with respect to the number of pixels for same features that are classified from supervised classification technique. The first row is for the classified features from supervised classification and they are considered as reference classification data or database. The first column is for extracted features from the proposed techniques that is investigated through this research. The pixel based comparative study is applied by comparing the number of classified pixels for same features from the used classification and extraction techniques.

Table 1 Confusion matrix for classification results

Results	Reference classification			Total
	Buildings	Roads	Vegetation	
Buildings	237920	0	610	238530
Roads	72025	49116	3406	124547
Vegetation	10178	34	49112	59324
Total	320123	49150	53128	422401
Percentage	74.3%	99.9%	92.4%	79.6%

The proposed extraction technique succeeded in extracting buildings with 74.3% successful percentage and extracting roads with 99.9% and vegetation areas with 92.4% successful percentages.

In general, the proposed extraction technique succeeded in extracting urban area features with 79.6% successful percentage.

VII. CONCLUSIONS

The proposed technique is considered as a pixel based urban area feature extraction technique. The proposed technique uses a combination from segmentation and classification methods to produce an efficient extraction of urban area features. *RGB* Color channels are used for identifying vegetation and shadows areas. Color image has to be transformed from RGB color space to *YCbCr* and *HSI* color spaces to luminance, hue and saturation color channels that are used for extracting roads and buildings. The proposed technique succeeded in identifying color ranges representing unhealthy vegetation areas.

The proposed feature extraction technique succeeded in extracting urban area features (buildings, roads and vegetation areas) with approximately 80% successful percentage, so it is possible to evaluate the proposed technique is fully automated and suitable for all urban areas using high resolution RGB images.

It has some difficulties in case of existing similarity in texture properties of different features. This problem appears, for example, in case of building with cement roof tops and roads that are constructed from concrete. So it is recommended to use additional information such as height properties to be able to extract, separate, and classify different features efficiently.

Extraction of different features shared texture properties without need of spatial information is a challenge. This challenge is as a motivation for continue this research by investigate a proposed technique for extraction all urban area features using color channels other than RGB from multispectral images instead of RGB image.

ACKNOWLEDGMENT

I acknowledge the Military Technical College (MTC), Egyptian armed forces (EAF) and my colleges in civil engineering department for continue support to complete and develop my researches.

REFERENCES

- [1] Ghanma, M. 2006. Integration of photogrammetry and LIDAR. Canada: University of Calgary (Canada), p. 156.
- [2] Tuncer, O. 2007. Fully automatic road network extraction from satellite images, Piscataway, NJ, USA, p. 708-14.
- [3] Sirmacek, B, and Unsalan, C. 2008. Building detection from aerial images using invariant color features and shadow information. 23rd International Symposium on Computer and Information Sciences, 2008. ISCIS '08. , p. 1-5.
- [4] Shorter, N, and Kasparis, T. 2009 Automatic Vegetation Identification and Building Detection from a Single Nadir Aerial Image. Remote Sensing;1(4):731-57.
- [5] Song, Y, and Shan, J. 2008 Buiding Extraction from High Resolution Color Imagery Based on Edge Flow Driven Active Contour and JSEG. Remote Sensing and Spatial Information Sciences;XXXVII(Part B3a):185 - 90.

- [6] Bong, DBL, Lai, KC, and Joseph, A. 2009 Automatic Road Network Recognition and Extraction for Urban Planning. Proceedings of World Academy of Science: Engineering & Technology;53:209-15.
- [7] Gonzalez, RC, Woods, RE, and Eddins, SL. 2004 Digital Image processing using MATLAB Pearson Prentice Hall.
- [8] Otsu, N. 1979 A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics;9(1):62-66.
- [9] Gonzalez, RC, and Woods, RE. 2002 Digital image processing Prentice Hall.
- [10] Boyer, KL, and Unsalan, C. 2005 A system to detect houses and residential street networks in multispectral satellite images. Computer Vision and Image Understanding;98(3):423-61.
- [11] Dance, C, Willamowski, J, Fan, L, Bray, C, and Csurka, G. 2004. Visual categorization with bags of keypoints. ECCV International Workshop on Statistical Learning in Computer Vision.