Comparative Study of Landmark Detection Techniques for Airport Visibility Estimation

J-Philippe Andreu[1], Harald Ganster[1], Erich Schmidt[2], Martina Uray[1], Heinz Mayer[1]

[1]Institute of Information and Communication Technologies
Joanneum Research, Austria
{jean-philippe.andreu, harald.ganster, martina.uray, heinz.mayer}@joanneum.at

[2]Austro Control
Österreichische Gesellschaft für Zivilluftfahrt mit beschränkter Haftung, Austria
Erich.Schmidt@austrocontrol.at

Abstract

Reliable and exact assessment of visibility is essential for safe air traffic. In order to overcome the drawbacks of the currently subjective reports from human observers, we present an approach to automatically derive visibility measures by means of image processing. It is based on identification of visibility of individual landmarks and compiling an overall visibility range. The methods used are based on concepts of illumination compensation as well as structural (edges) and texture recognition. Validation on individual landmarks showed a reliable performance of 96% correct detections. Furthermore, a solution for compiling the overall visibility report is presented, that resembles the currently used standard in air traffic management.

1. Motivation

In order to guarantee for safe air traffic controllers are relying on precise forecasts and measurements of the current weather situation. The exact acquisition and specification of the atmospheric condition builds the basis for any forecast and thus is elementary for any aviation weather service. While weather minima[1] that still allow for efficient and safe air traffic are continuously lowered, they still have huge impact on air traffic. E.g. thunderstorms [13, 2] along with all their weather phenomena have severe influence on efficient and economic handling of air traffic. This is also reflected in various delay-statistics [1, 13].

Figure 1: Classic airport sensor system: meteorological instrumentation and equipments (left), ceilometer (middle), sketch of a visibility sensor (right).

[1] Visual meteorological conditions: http://en.wikipedia.org/wiki/Visual_meteorological_conditions

All major airports operate dedicated sensor systems to assess the current weather situation. Besides "classic" parameters like wind, pressure, humidity, and temperature, there are point-like measurements of visibility and cloud cover information (Figure 1).
One of the essential parameters is a precise measurement of the visibility in the airport range. Currently human observers compile visibility reports[2] every 30 minutes based on visual observation of known landmarks (prominent structures like buildings, mountain tops, etc...) in a so-called landmark map (Figure 5 and Figure 6). These landmarks have attached the distances to the observation point, and by identifying which of those landmarks are still visible, and which not, a visibility estimate is derived.
This naturally is very subjective to the individual observer and error-prone, thus an objective measurement is highly sought by all operators to allow for efficient flight and tactical air traffic planning as well as for operational handling of air traffic. The automated measurement with a visibiliy sensor (Figure1, right) is still a rough estimate only as it calculates the overall long-range visibility from close and very local observations.
The approach presented in this paper is aiming at emulating the human observer procedure, by deriving the visibility of already established landmarks with automatic image processing methods. As the landmarks completely differ from each other in their appearance, structure, size, as well as by illumination variations (time of the day, changing meteorological conditions such like rain, snow and fog), it is not possible to set-up a unique method that can cope for all landmarks. Thus, different approaches to illumination compensation, structure detection, and classification are applied and evaluated, in order to find the most discriminating method for each individual landmark.
The rest of the paper is structured as follows. The methods applied to landmark detection are presented in Section 3 following a short overview of the relevant literature in this field (Section 2). By analysis of a comparative study (Section 4) it is shown how to choose, for each landmark separately, the best suited method for its recognition. This is followed by displaying the accuracy of the method on several landmarks and a discussion on the strategy in assessing the prevailing visibility based on the visibility of each singular landmark. Section 5 concludes with final remarks.

2. Related Work

Current visibility sensors employ the sender and receiver principle: a ray of light is emitted by a projector and caught either by a photodetector (e.g. scatter meter) or by a digital camera. The literature review below shows that the usage of cameras for the measurement of visibility became more and more popular over the last six years.
In 2005, Luo et al. [9] measured visibility by analyzing the intensities of grey level images. Due to the fact that high-frequency information depends on the brightness and the texture in urban images they developed a model to establish a relationship between frequency components and urban visibility. They showed that using a Sobel operator or FFT (high pass filtering) is adequate for extracting high frequency components and thus for monitoring visibility. Furthermore it was proven that the results of both methods correlate with each other as well as with human observations.
Also Raina et al. [11] investigated the usage of contrast for the measurement of visibility. Unlike [9], instead of investigating the whole image, only regions of interest were employed. Their experiments are based on a network of webcams where contrast values of acquired images are compared to clean day conditions. Statistical evaluations finally allow for the sought classification.

[2] METAR: international standard format for reporting weather information

Another approach was presented by Kim et al. [7] who investigated the relationship between the optical measurement and HSI colour differences. Their goal was to analyze air pollution based on visibility. The idea for the approach itself is based on the fact that the colour of sky depends on the light scattering (e.g. blue for small aerosol particles and white for larger particles) and especially the colour of haze varies with the optical properties of aerosol. By measuring the difference of the HSI space between a target image and the clear sky reference image it is possible to estimate the status of visibility by the usage of the developed grading visibility level.
Poduri et al. [10] went one step further trying to make sky analysis available for mobile phones. Their approach is based on the generation of an analytic model of the sky as a function of appearance. Visibility is finally estimated by comparison of a new image with this model. The main drawback of this method is that it works for cloud free sky only.
All those approaches have in common that their main goal is to "see" the amount of pollution in the air. The basic idea is always to develop a reference model and compare it to the newly acquired image. In contrast to our work we are not interested in the long distance visibility only but especially in the maximum visible distance represented by the visibility of previously defined regions of interest. Furthermore, we do not use a "perfect weather image" as reference image but employ several occurring views.

3. Landmark Detection

Our visibility estimation is based on recognition of ground landmarks scattered around airports. Due to different sun illumination throughout the day and varying weather conditions like rain, snow and fog, the same landmark can display drastically different appearances. Several approaches have been proposed for solving the variable illumination problem in image processing. As shown in [12] for face recognition under various lighting conditions these approaches can be classified into three main categories: *normalization*, *invariant features extraction* and *modelling*.
The first category of approaches (*normalization*) includes image pre-processing algorithms that are employed to compensate and normalize the illumination. Since most of these algorithms do not require any training or modelling they can be considered as general purpose image pre-processing algorithms like histogram equalization, gamma correction, and logarithmic transforms.
The second category (*invariant features extraction*) aims at extracting illumination invariant features from the image and applies the recognition on those. Edge maps and different texture descriptors (Gabor filtering, Local Binary Patterns, etc...) belong to this category.
The last category of approaches (*modelling*) generally uses low-dimensional linear subspaces for modelling image variations under different lighting conditions. For instance, the Principal Component Analysis (PCA) falls among this category. These approaches generally require a training set of images representing the object under a lot of different illumination conditions.

3.1. Illumination normalization

The Retinex theory [8] yields an algorithm for extracting an illumination-normalized representation of the image. The theory is based on the reflectance illumination model of human vision which assumes that it is sensitive to scene reflectance and local change of contrast while being insensitive to illumination conditions and global brightness levels. The basic definition for a pixel at position (x, y) is:

$$R(x,y) = \log I(x,y) - \log\left(Ke^{-r^2/c} \otimes I(x,y)\right) \tag{1}$$

with $K = 1/2\pi\sigma^2$, $r^2 = x^2 + y^2$ and $c = 2\sigma^2$

This is the logarithmic difference (e.g. the logarithm of the quotient) between the image and a version of it convolved with a low-pass filter (i.e. Gaussian filter), the low frequency component being considered as the illuminance of the input stimulus (Figure 2).

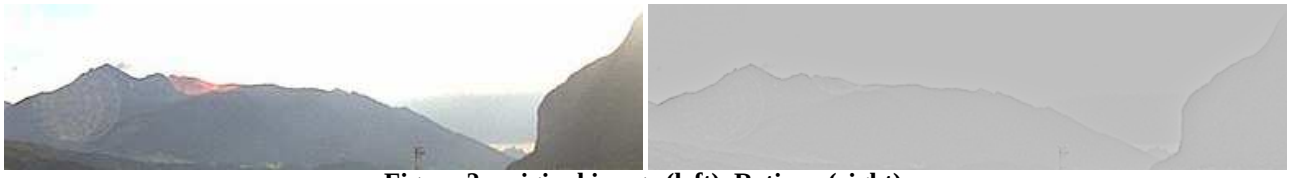

Figure 2: original image (left), Retinex (right)

The constant *c* is referred as the scale of the Retinex and by varying it, Jobson et al. [6] proposed an extension to the original Retinex algorithm that instead of giving the result at a single scale, outputs an image constructed as the weighted sum of single scaled Retinex images. That implementation displays a better balance of dynamic compression and colour rendition.

3.2. Invariant features

Edges are one of the main illumination invariant features in images. The Canny edge detector [4] is a well-known edge detection algorithm adaptable to various environments. Its parameters for recognition of edges of differing characteristics depending on the particular requirements of a given implementation. It uses a multi-stage algorithm to detect a wide range of edges in images. It is aimed at discovering the "optimal" edge detection algorithm: optimal in the way that the algorithm should find as many real edges in the image as possible, the edges marked should be spatially as close as possible to the real edges and a given edge in the image should not be doubled, and possibly, image noise should not create false edges (Figure 3).

Figure 3: original image (left), Canny edge detector (right)

Another type of illumination invariant features is the distribution of local intensity gradients or edge directions. Local object appearance and shape can often be characterized by this kind of distribution, even without precise knowledge of the corresponding gradient or edge positions. That can be implemented by dividing the image window into small spatial regions (e.g. "cells"), for each cell accumulating a local 1-D histogram of gradient directions or edge orientations over the pixels of the cell. For better invariance to illumination, it is also useful to contrast-normalize the local cells over somewhat larger spatial regions (i.e. "blocks"). These normalized descriptor blocks are referred as Histogram of Oriented Gradient (HOG) descriptors [5]. Operatively an image is tiled with a dense (e.g. overlapping) grid of HOG descriptors and the resulting combined feature vector is then used for achieving recognition.

3.3. Modelling

Numerous computer vision systems employ the appearance-based paradigm for object recognition. One primary advantage of appearance-based methods is that it is not necessary to create representations or models for objects, since, for a given object its model is implicitly defined by the selection of the sample images of the object. The most popular technique is the Principal Components Analysis (PCA) which was first used for face recognition [15], and is herein used on different appearances of landmarks.

4. Comparative study

The presented study compares the different selected methods of illumination compensation by applying them on a large data set of real images of landmarks. In order to compare the methods a common measure is used. Based on that measure we show how to select the one giving the best result for detecting a landmark. For two representative landmarks, we assess the accuracy of the selected method. We end-up by showing how to get the prevailing visibility based on the visibility of each separate landmark.

4.1. Training and test data

All tests have been carried out on images of landmarks from two different airports: Innsbruck (Austria) and Graz (Austria). From the first airport 15 landmarks (Figure 5) were used to assess the visibility distance while 35 were used from the second one (Figure 6).

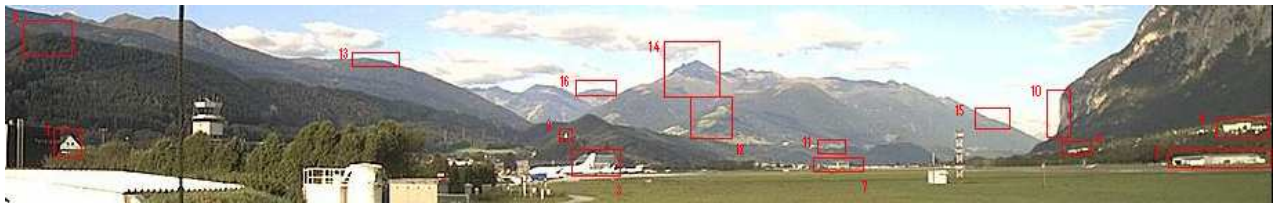

Figure 5: Innsbruck airport landmarks

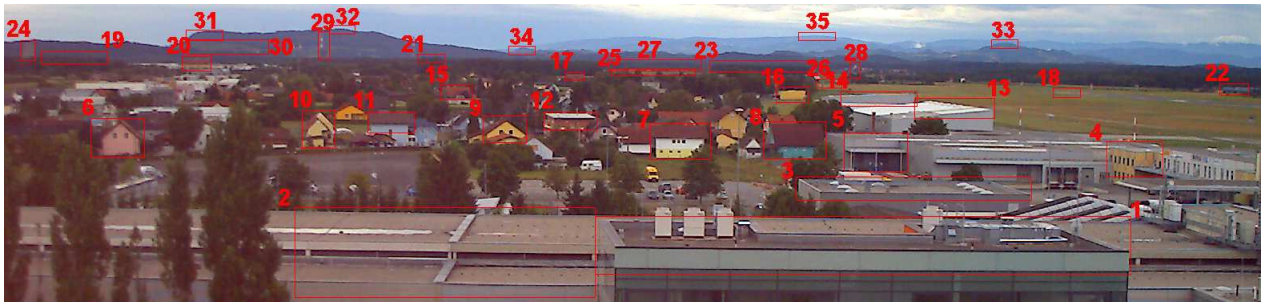

Figure 6: Graz airport landmarks

For each airport and each given landmark three sets of images were derived: a training set consisting of images with visible landmark, a test set with visible landmark, and a test set where the landmark was *not* visible. All sets were equally sized and all together 4500 images were selected for our experiments. Note that we also considered using training sets consisting of images where the landmark were *not* visible but experimental evaluation showed that this approach yielded worse recognition results than building training sets of images with visible landmarks. This behaviour can be explained by the fact that images where the landmark is not visible shows far less appearance variability than visible landmark images, thus lowering the discriminative power of the training set.

4.2. Method comparison

Instead of evaluating the different illumination compensation methods solely against each other, we inserted a method for landmark detection not compensating for illumination in the comparison: simple template matching [3] with a noise removal but edge preserving preprocessing: bilateral filtering [14]. This method is non-iterative, local, simple, and smoothes images while preserving edges. It is designed, by means of a nonlinear combination of domain and range filtering, to prevent averaging across edges while smoothing an image. It combines gray levels based on both their geometric closeness and their photometric similarity, and prefers near values to distant values in both domain and range (Figure 7).

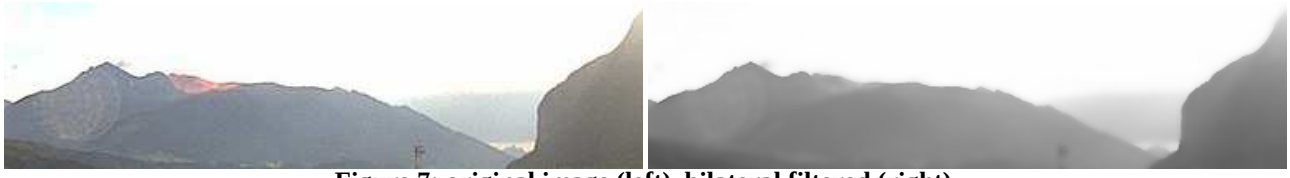

Figure 7: original image (left), bilateral filtered (right)

In order to compare the results of the different methods on the different test sets, a common performance measure is necessary. Classically in image processing the cross-correlation is used for the similarity estimation of two images. It is expressed mathematically as follows:

$$\frac{1}{n-1}\sum_{x,y}\frac{(f(x,y)-\overline{f})(t(x,y)-\overline{t})}{\sigma_f\sigma_t} \qquad (2)$$

This measure is appealing since it can easily be normalized between 1 (meaning the two images are identical) and -1 (meaning the two images are opposite). While it is easy to cross-correlate the images resulting from applying the Retinex, Canny edge detection and bilateral filtering methods, it is a less straightforward task for the appearance-based paradigm method (e.g. PCA). Note that for the HOG method, even though its output is not an image, the feature vectors (HOG descriptors) can directly be compared with cross-correlation. PCA is building, with a dimensional reduction, an eigenspace out of an image training set. Projecting test images into that space and computing their distance to the nearest projection of a training image cannot be normalized. To alleviate that problem, we use the reconstruction (e.g. back-projection or approximation) of the test image from its eigenspace coordinates for cross-correlation purposes.

4.3. Method selection

In order to select the method best suited for detecting a given landmark, we first evaluated for each method the test set containing the positive samples (images where the landmark is visible) and the test set containing the negative samples (images where the landmark is *not* visible) against their respective training set. Then, for each method, we computed the mean and standard deviation of the performance of the two test sets. One can easily figure out that the performance on the test set containing the positive samples should be higher than the one on the test set containing the negative samples. In order to quantify the discrimination power we assume that the bigger the interval between the mean (minus its standard deviation) of the performance on the test set containing the positive samples and the mean (plus its standard deviation) of the performance on the test set containing the negative samples, the better the method is suited for detecting the landmark under investigation. The two tables in Figure 8 and Figure 9 show highlighted in green which method scored best for each individual landmark. The nature of the landmark itself (e.g. textured, edge based, colored) influences the automatic selection of its appropriate detection method. For instance, structural landmarks (buildings, mountain tops) tend to favour the edge-based approaches, whereas plain landmarks (meadows, forests) are better detected by textural algorithms.

Land Mark	Retinex Min [ ]	Canny Min [ ]	Bilateral Min [ ]	PCA+ Min [ ]	HOG Min [ ]
1	0.129	0.065	**0.198**	0.140	0.190
2	0.189	0.048	0.279	**0.406**	0.263
3	0.151	0.047	0.335	**0.345**	0.220
4	0.330	0.205	**0.339**	0.159	0.283
5	**0.315**	0.267	-0.021	0.083	0.206
6	0.021	**0.237**	-0.010	-0.039	0.111
7	**0.168**	0.107	-0.112	-0.170	0.143
9	0.111	0.065	0.200	**0.292**	**0.292**
10	**0.296**	0.097	0.023	0.028	0.166
11	**0.145**	0.065	0.068	-0.253	0.034
12	0.247	0.083	0.191	-0.097	**0.353**
13	0.331	0.196	0.064	0.184	**0.404**
14	**0.262**	0.074	-0.336	-0.071	0.203
15	0.439	0.259	0.250	0.124	**0.521**
16	**0.200**	0.115	-0.135	0.062	0.116

Figure 8: Landmark method selection (Innsbruck airport)

Land Mark	Retinex Min [ ]	Canny Min [ ]	Bilateral Min [ ]	HOG Min [ ]	PCA+ Min [ ]
1	0.283	0.113	0.149	**0.516**	0.427
2	0.296	0.085	0.363	**0.457**	0.424
3	0.128	0.046	0.341	0.400	**0.618**
4	0.131	0.060	-0.134	**0.530**	0.313
5	0.085	0.065	-0.122	0.434	**0.608**
6	0.496	0.189	0.624	0.580	**0.708**
7	0.239	0.192	0.636	0.501	**0.726**
8	0.182	0.160	0.552	0.358	**0.707**
9	0.360	0.122	0.371	0.581	**0.598**
10	0.249	0.149	0.581	0.496	**0.724**
11	0.373	0.148	0.070	0.573	**0.651**
12	0.030	0.067	0.156	0.386	**0.570**
13	0.233	0.070	**0.495**	0.431	0.425
14	0.099	0.028	0.187	**0.538**	0.379
15	0.159	0.069	0.434	0.379	**0.563**
16	0.168	-0.036	0.420	**0.615**	0.528
17	-0.026	0.000	0.003	**0.306**	0.065
18	0.166	0.000	0.417	**0.438**	0.106
19	0.469	0.028	0.120	**0.468**	0.400
20	0.229	-0.009	0.099	0.294	**0.447**
21	0.210	0.116	0.089	**0.472**	0.229
22	0.118	-0.007	0.342	**0.409**	0.229
23	0.319	0.039	0.158	**0.350**	0.173
24	0.210	-0.011	-0.012	0.235	**0.443**
25	-0.190	-0.210	-0.140	**0.009**	-0.015
26	0.045	-0.079	**0.145**	0.103	-0.098
27	0.085	-0.294	0.017	**0.207**	0.173
28	-0.269	-0.018	-0.146	-0.071	**0.035**
29	0.224	0.196	-0.188	**0.251**	0.144
30	0.149	0.000	-0.185	**0.150**	0.015
31	0.381	0.133	0.275	0.327	**0.434**
32	0.112	0.055	0.034	0.458	**0.489**
33	0.350	0.000	0.253	0.244	**0.371**
34	0.172	0.108	0.060	0.234	**0.346**
35	0.224	0.000	0.187	**0.476**	0.337

Figure 9: Landmark method selection (Graz airport)

4.4. Detection accuracy

Estimating the accuracy of individual landmark detection for a selected method is a very time consuming task since there is no ground truth data. Every image has to be visually verified by a human operator for the visibility (or not) of a specific landmark.

We performed that task for the landmark #12 (a meadow situated roughly 10km away from the control camera) of the Innsbruck airport and evaluated on more than 15000 images how the selected method (in this case HOG) is performing. The recognition rate was about 96% and among the 4% misses the rate of false positives and false negative was equally spread (roughly 2% each).

We did the same for the landmark #22 (an airport maintenance building situated 2200m away from the control camera) of the Graz airport and evaluated on more than 4000 images how the selected method (HOG again) is performing. The recognition rate was again about 96% and among the 4% misses with a rate of 3% for false positives and 1% for false negatives.

4.5. Visibility estimation

Finally a global visibility estimate is inferred from the visibility of each individual landmark. We know for each individual landmark its distance to the control camera thus we can order them by their corresponding distance. For the Innsbruck airport the landmarks are spread from a distance of 250m to a distance of 20km while for the Graz airport, the landmark distances range from 90m to 55km. There are two natural ways to automatically infer the global visibility: either set the visibility to the distance of the furthest visible landmark, not caring if in between one landmark was not detected (optimistic approach) or set the visibility to the distance of the last detected landmark in the uninterrupted sequence of visible landmarks (conservative approach). While the first approach is more sensitive to false positives (e.g. a landmark is reported as visible while it is not), tests showed that this approach delivers distance estimation results which are comparable with the visibility reported by current METAR reports.

5. Conclusion

Variable illumination is a major problem in landmark detection and recognition. The aim of this work was to investigate several illumination compensation and normalization algorithms. By the presented comparative study, five different image processing methods were tested on a total of 4500 images. The methods were compared and a measure for selecting the most appropriate method for each individual landmark was discussed. The accuracy of the selected methods was visually

assessed for two landmarks on more than 15000 images and it displayed a very satisfying recognition rate of 96%. Finally, we proposed a method to infer the overall airport visibility from individual landmark visibility. Setting the overall visibility distance to the distance of the furthest visible landmark resembled results to what is reported by air controllers. Anyway, having the knowledge of each individual landmark visibility opens the field for a new quality in visibility reports (e.g. visibility in several directions, locations) compared to the currently used single distance measures only.

6. Acknowledgments

This research was supported by the TAKE OFF programme, an initiative of the Austrian "Federal Ministry for Transport, Innovation and Technology" under contract number 820742.

7. References

[1] AEA – the Association of European Airlines: AEA CONSUMER REPORT FOR 4th QUARTER AND ANNUAL 2007. http://files.aea.be/News/PR/Pr08-006.pdf

[2] ALLAN, S.S., BEESLEY, J.A., EVANS, J.E. and GADDY, S.G., Analysis of Delay Causality at Newark International Airport, in: 4th USA/Europe Air Traffic Management R&D Seminar, (2001), Santa Fe, USA.

[3] BRUNELLI, R., A Template Matching Techniques in Computer Vision: Theory and Practice, Wiley, ISBN 978-0-470-51706-2, 2009.

[4] CANNY, J., A Computational Approach To Edge Detection, in: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6 (1986), 679-698.

[5] DALAL, N. and TRIGGS, B., Histograms of Oriented Gradients for Human Detection, in: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1 (2005), 886-893.

[6] JOBSON, D.J., RAHMAN, Z. and WOODELL, G.A., A multiscale retinex for bridging the gap between color images and the human observation of scenes, in: IEEE Trans. on Image Processing, Vol. 6, No. 7 (1997), 965-976.

[7] KIM, K.W. and KIM, Y. J., Perceived Visibility Measurement Using the HSI Color Difference Model, in: Journal of the Korean Physical Society, Vol. 46, No. 5 (2005), 1243 – 1250.

[8] LAND, E.H., The retinex theory of color vision, in Scientific American, Vol. 237, No. 6 (1977), 108–128.

[9] LUO, C., WEN, C., YUAN, C., LIAW, J., LO, C. and CHIU, S., Investigation of urban atmospheric visibility by high-frequency extraction: Model development and field test, in: Atmospheric Environment, Vol. 39 (2005), 2245-2552.

[10] PODURI, S., NIMKAR, A. and SUKHATME, G. S., Visibility Monitoring using Mobile Phones, http://robotics.usc.edu/~mobilesensing/visibility/MobileAirQualitySensing.pdf, (2010).

[11] RAINA, D. S., PARKS, N. J., LI, W., GRAY, R. W. and DATTNER, S. L., An Innovative Methodology for Analyzing Digital Visibility Images in an Urban Environment, in: Journal of Air and Waste Manage, Vol. 55 (2005), 1733-1742.

[12] RUIZ-DEL-SOLAR, J. and QUINTEROS J., Illumination compensation and normalization in eigenspace-based face recognition: A comparative study of different pre-processing approaches, in: Pattern Recognition Letters, Vol. 29 (2008), 1966-1979.

[13] SASSE, M. and HAUF, T., A study of thunderstorm-induced delays at Frankfurt Airport, Germany, in: Meteorological Applications, Vol. 10 (2003), 21-30.

[14] TOMASI, C. and MANDUCHI, R., Bilateral Filtering for Gray and Color Images, in: Proceedings of the IEEE International Conference on Computer Vision, (1998), Bombay, India.

[15] TURK, M. and PENTLAND, A., Eigenfaces for Recognition, in: Journal of Cognitive Neuroscience, Vol.3, No.1 (1991), 71-86.