3D Visualization via Augmented Reality: The Case of the Middle Stoa in the Ancient Agora of Athens

Styliani Verykokou[1], Charalabos Ioannidis[1] and Georgia Kontogianni[1]

[1] Laboratory of Photogrammetry, School of Rural & Surveying Engineering,
National Technical University of Athens, Greece
st.verykokou@gmail.com, cioannid@survey.ntua.gr, georgia.kontog@gmail.com

Abstract. Augmented reality is a rapidly evolving technology that enriches reality with computer generated information as well as a powerful tool that provides innovative ways of information access at cultural heritage sites. In this paper, an augmented reality application that allows the visualization of a part of the Middle Stoa in the Ancient Agora of Athens is presented. Users of this application, pointing their tablet PC at the present situation, have the opportunity to see what this building looked like in ancient times, as its three dimensional model is displayed on the camera view of their device, projected on the modern-day ruins.

Keywords: Augmented Reality, Cultural Heritage, Photogrammetry, Computer Vision, 3D model

1 Introduction

Augmented reality (AR) is the scientific field that deals with the combination of the real world with computer generated data. It aims at the integration of synthetic information in the live view of the physical world, which is not entirely hidden but has the predominant role instead, and enriches people's perception of reality. AR belongs to the technology of mixed reality [1], according to which objects that belong to both the real and virtual world are presented as coexisting in the same place.

Although AR was first presented in the 1960s with the creation of a head-mounted three dimensional (3D) display [2], it was coined much later, in 1992 [3], and since then it has begun to have practical applications in various fields, including culture, archaeology, tourism, sports, entertainment, architecture, art, the army, education, medicine, advertising, navigation, commerce, interior design and task support [4]. AR systems combine real and virtual objects in a real environment, allow real time interaction and register virtual objects in the 3D space [5]. Their basic components include a display (i.e., head-worn, handheld or spatial display), a computer system, a camera or another optical instrument, appropriate software, the real scene and the virtual objects as well as – depending on the application – various sensors, markers or patterns for recognition, web services, a content server and possibly more [6], [7].

During the past decade, AR has offered interesting possibilities and useful applications for the disclosure of cultural heritage, the promotion of historic materials and the interactive visualization of heritage items [8]. One of the first such AR systems, named Archeoguide, was built around the archaeological site of Olympia in Greece during the period from 2000 to 2002 [9]. Visitors equipped with a head-mounted display (HMD) with an external camera and a compass, as well as a backpack with a computer, a GPS receiver, a battery and wireless communication equipment had the ability to see 3D monuments as they were in antiquity, rather than the present ruins, and virtual athletes competing in the stadium. Alternatively, a pen-tablet PC or a palmtop computer combined with proper equipment could be used, although they did not provide users with the full AR experience.

Another project that took place during 2002-2004 is LIFEPLUS. Its aims were the revival of life in frescos of ancient Pompeii through the real time rendering of realistic 3D simulations of animated virtual humans re-enacting staged storytelling dramas, animals and plants, in front of visitors equipped with an HMD, earphone and mobile computing equipment, as well as the provision of continuous guidance to visitors, including the addition of historic audio-visual information [10]. In the same period, a tablet was proposed to be used as AR device in Els Vilars, an archaeological site in Catalonia, in order to provide visitors with cultural and educational multimedia content, to wit reconstruction models and additional information [11].

The iTacitus project, which was completed during the period between 2006 and 2009, aimed at the encouragement of cultural tourism. One of the systems developed under this project was the AR presentation system for remote cultural heritage, which rendered additional information on top of the camera feed of an Ultra Mobile PC (UMPC) or a stationary AR see through device, named MovableScreen. The UMPC was used for the superimposition of digital grain plans on a satellite image of Berlin's center, in order to illustrate its urban development from 1940 to 2008 and the course of the Berlin Wall, by showing how its 3D model changed during that period. Furthermore, visitors of the exhibition area of SIGGRAPH 2008 pointing the UMPC camera at a floor map of the ancient Forum Romanum were able to see 3D models of monuments of Ancient Rome, constructed for the "Rome Reborn" project, superimposed on the map. The UMPC and the MovableScreen were also used in an exhibition of the Allard Pierson Museum in Amsterdam in two AR applications about Satricum and Forum Romanum. According to them, visitors had the ability to see supplementary information (i.e., 3D models, photographs, etc.), while pointing these devices at two large photographs on the walls of the exhibition space [12].

Another approach of using AR in museums was adopted in 2008 by the DNP-Louvre Museum Lab in Tokyo, where visitors equipped with a portable device with a camera could see on the live feed a virtual character giving them details about the visit, as well as 3D models and additional information [13]. Moreover, both screen-based and helmet-based AR were used in the Boijmans van Beuningen Museum in Rotterdam in the period 2008-2009, for the visualization of cultural heritage and the addition of audiovisual information [14]. The interest towards AR technology shown by museums was continued in subsequent years. In 2009, the Powerhouse Museum in Sydney and in 2010, the Museum of London released two AR applications for mobile platforms that display historic photographs of Sydney and London respectively, combined with the appropriate description [15], [16].

In 2010, the mobile application CultureClic was launched in France. It allows visitors to discover geolocated works of art, see what the place where they are located looked like several centuries ago, access information about museums and discover cultural events in AR mode [17]. In the same year, the Netherlands Architecture Institute launched a mobile application that, inter alia, allows people to see how several places in the Netherlands used to be in the past [18]; the City of Philadelphia Department of Records started developing an AR application that would enable users to see historic photographs of Philadelphia as overlays on the camera view of their smartphones [19]. Additional research into combining historic images and information with AR in mobile phones was undertaken in 2010 in San Francisco [20] as well as in 2012 in a historical street in Norway [21]. A recent project that started in 2013 and is not yet completed is TAG CLOUD, which aims to amplify the active participation of general public in cultural heritage activities via social media, AR and storytelling applications based on the cloud [22].

2 Application Development

A pattern based markerless AR application that augments the real scene viewed on the camera of a mobile device with a 3D model was developed. In this section, the methodology followed and the implementation of the application are presented.

2.1 Methodology

The application is based on the recognition of an almost planar object. The initial data that is demanded includes the interior orientation of the camera that captures the real world (i.e., the pixel coordinates of the principal point (x_0, y_0), the camera constant in pixels in x and y direction (c_x, c_y) and the coefficients of lens distortion polynomials), a pattern image, which is an image of the object which has to be recognized, and the 3D model that augments the real scene. A simplified flowchart representing the methodology that is described in this subsection is illustrated in Fig. 1.

Definition of the Coordinates of the Corners of the Pattern Object. The origin of the object coordinate system is located at the center of the pattern object, X and Y axes lie on the object plane, while Z axis is perpendicular to it. X and Y coordinates of the four corners of the pattern object derive from the normalized width and height of the pattern image, thus ranging from -1 to 1, whereas Z coordinate is set to zero.

Features Detection and Description. Features extraction is applied in the pattern image and in every real world frame, after they are converted to greyscale, using the SURF (Speeded-Up Robust Features) algorithm [23]. Its first step is the detection of feature points located in blob-like structures of an image, based on the determinant of an approximation of the Hessian Matrix, which is computed for every pixel of the image in all scale levels of each octave into which scale space is divided. The second step is the computation of a 64-dimensional vector for each interest point, called descriptor, which indicates the underlying intensity structure of a square region around it, oriented along its dominant orientation, with the size of the region depending on the scale of the point. The extracted feature points are scale and rotation invariant, while skew, anisotropic scaling and perspective effects are also covered to some degree.

Matching. The matching stage is executed through the process of finding correspondences between the pattern image and every real world frame. For each interest point in every frame, the two nearest feature points of the pattern image are detected, based on the Euclidean distance between their descriptors. If the Nearest Neighbor Distance Ratio (NNDR), which is computed according to equation (1), is smaller than a threshold and if the distance between the descriptors of the feature point in the real world frame and the nearest point in the pattern image is smaller than a maximum accepted value, this correspondence is considered to be valid.

$$\mathrm{NNDR} = \frac{\|v - v_1\|}{\|v - v_2\|} \,, \qquad (1)$$

where: v is the descriptor of an interest point in the real world frame,
v_1 and v_2 are the descriptors of the first and the second nearest feature point, respectively, in the pattern image.

However, many incorrect matches remain. These outliers are rejected via the RANSAC (RANdom SAmple Consensus) algorithm [24], if a minimum number of five matches is detected. Otherwise, the scene will not be augmented, because it is considered that the pattern object cannot be recognized in the frame. RANSAC calculates the geometric relation between the pattern image and the image of the pattern object in each frame using the detected correspondences, relying on the use of the minimum number of data. This relation is considered to be the homography in the two dimensional projective space and is calculated using the best sample of four matches. The correspondences that verify the computed homography are detected and constitute the inliers. If at least five inliers are detected, the scene will be augmented.

Homography Estimation and Pattern Recognition. The homography that relates every point of the pattern object depicted in a real world frame with the corresponding point of the pattern image is expressed by an invertible 3x3 matrix, with the use of homogeneous coordinates (equation (2)).

$$\begin{bmatrix} x_{frame} \\ y_{frame} \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{pattern} \\ y_{pattern} \\ 1 \end{bmatrix} , \qquad (2)$$

where: x_{frame}, y_{frame} are the pixel coordinates of a point in a frame,
$x_{pattern}$, $y_{pattern}$ are the pixel coordinates of the same point in the pattern image,
h_{ij} are the elements of the homography matrix.

The initial estimation of the homography obtained by RANSAC is refined using all the inliers, via the Levenberg-Marquardt nonlinear optimization algorithm [25], [26], in order to minimize the reprojection error.

The recognition of the pattern object in each real world frame is accomplished through the calculation of the pixel coordinates of the four corners of the pattern object. These are computed according to equation (2), using the estimated homography matrix and the pixel coordinates of the four corners of the pattern image.

Exterior Orientation Estimation. The camera exterior orientation for every frame is computed using the interior orientation parameters, the pixel coordinates of the corners of the pattern object in the frame and their corresponding object coordinates. The mathematical model used is the projection transformation (equation (3)).

$$\lambda \cdot \begin{bmatrix} x_{frame} \\ y_{frame} \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} c_x & 0 & x_0 \\ 0 & c_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{K}} \cdot \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix}}_{[\mathbf{R}|\mathbf{t}]} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (3)$$

where: X, Y, Z are the object coordinates of a point x_{frame}, y_{frame} in the frame,
K is the matrix with the camera intrinsic parameters,
[**R**|**t**] is the joint rotation-translation matrix with the camera extrinsic parameters,
λ is a scale factor.

The elements of [**R**|**t**] are computed linearly after the undistortion of the image coordinates of the corners of the pattern object in the frame and the computation of the homography that relates the X, Y coordinates of the pattern object with the corresponding undistorted image coordinates [4]. The calculation of the singular value decomposition of the computed rotation matrix **R** follows in order to refine it, by coercing it to satisfy the orthogonality condition, as described in [27]. Afterwards, the rotation matrix is transformed to a 3D rotation vector, using the Rodrigues rotation formula [28]. Subsequently, the camera 6-DOF pose is optimized via the Levenberg-Marquardt algorithm and the rotation vector is converted back into a 3x3 rotation matrix using the Rodrigues formula. The outcome of this procedure is the joint rotation-translation matrix for each real world frame.

Rendering of the Augmented Scene. The initial step of this process is the rendering of the real world frame on a window, so that it forms its background. The last one is the rendering of the 3D model on that window. This procedure is described below.

The coordinates of the vertices of the 3D model in its local coordinate system are transformed to the object coordinate system by being normalized in the range [-1, 1], as well as being multiplied with the appropriate scale factor and with the proper translation and rotation matrices, so that the model is located at the right position, with the intended orientation and scale relative to the pattern object. These coordinates are transformed to the camera system, using the elements of [**R**|**t**] according to the viewing transformation [4]. The viewing volume of the camera is assumed to be a truncated pyramid, because of the used model of perspective projection. Camera coordinates are transformed to clip coordinates using the camera interior orientation parameters, the dimensions of the frames as well as the distances of the near and far clipping planes of the viewing volume from the projection center. Then, clip coordinates are converted into normalized Cartesian coordinates which are transformed to window coordinates in pixels by viewport transformation and a depth-range transformation is applied for the acquisition of depth information [4].

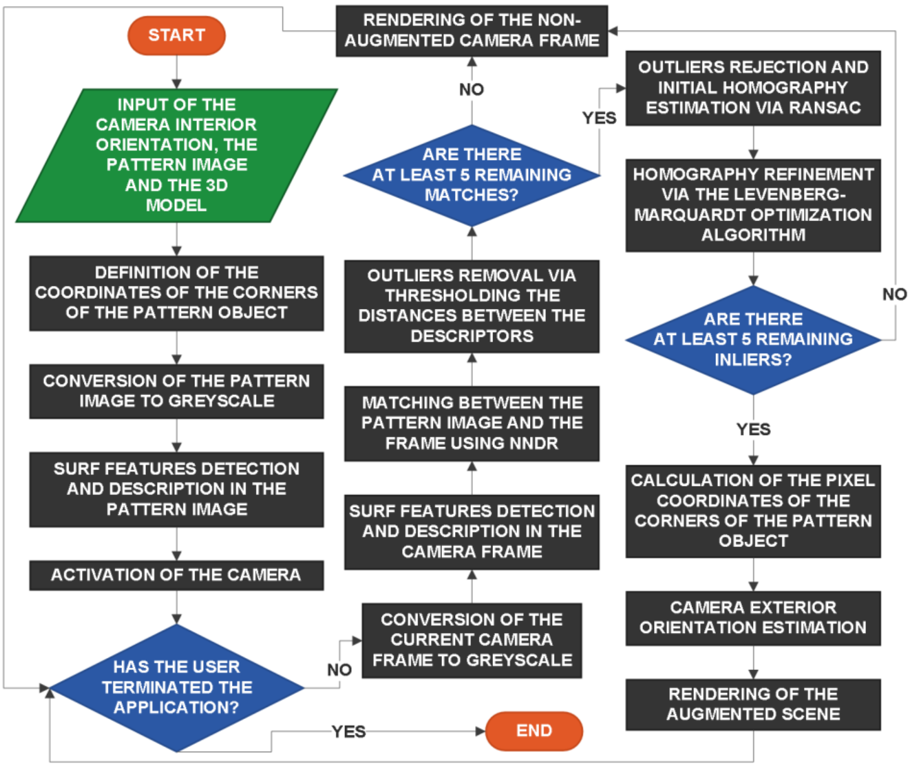

Fig. 1. Flowchart representing the methodology followed in the AR application.

2.2 Implementation

The application was developed in the C++ programming language. The OpenCV library was used for the procedure followed in order to calculate the camera exterior orientation for each real world frame. The OpenGL application programming interface was used for the rendering of the augmented scene, while library GLM: An Alias Wavefront OBJ file Library was used in order to load the 3D model.

The application is intended for tablet computers running Microsoft Windows; it can also be executed in any other computer with Windows operating system. It uses the tablet built-in camera in order to capture frames of the real world and augment them in almost real time and demands the camera intrinsic parameters as input. The interior orientation of the camera of the tablet PC was calculated using OpenCV code that was converted into executable file. The fully automated calibration, based on Zhang's [27] and Bouguet's [29] methods, was done by taking pictures of a planar chessboard shown at several different orientations.

3 Case Study

The aim of the developed AR application is the real time 3D visualization of the south side of the Middle Stoa in the Ancient Agora of Athens, Greece. It enables people to see what this building looked like in antiquity through an innovative AR experience. Thus, users of this application have the opportunity to see the ancient building on the camera view of their tablet in the same place where it was in the past as well as with the same dimensions and orientation.

3.1 The Middle Stoa in the Ancient Agora of Athens

The Ancient Agora of Athens, located to the northwest side of the Acropolis, is one of the most important archaeological sites in Athens. It was probably laid out in the 6^{th} century B.C. and it was the social, educational, administrative, philosophical, political, cultural, religious, commercial and financial center of the city [30]. The Middle Stoa was the largest building in the Agora, measuring about 147m x 17.5m, and divided it into two unequal sides, the north and the south one. It was presumably built between 180 and 140 B.C. and it was used for commercial purposes. It was a relatively modest building, made of limestone, with a terracotta roof. On all four sides it was enclosed by collonades, which consisted of 160 unfluted Doric columns, while in the middle 23 Ionic columns existed. Today only the foundations of the building and some individual parts of it are visible in the site [31], [32].

The 3D virtual reconstruction of a part of the Middle Stoa was performed in Autodesk 3ds Max software, according to reconstruction studies and drawings by the architect Travlos [33] and the American School of Classical Studies at Athens [34]. In cases where the information was not leading to firm conclusions, the advice and suggestions of specialized scientists responsible for the Agora proved to be valuable. A part of the south side as well as a part of the west side of the Middle Stoa were reconstructed. The 3D model comprises the krepis, the colonnade, the entablature, which consists of the architrave, the frieze and the cornice, the sima, the pediment and the inclined roof covered with ceramic tiles. As far as the colonnade is concerned, thirteen columns were reconstructed and nine of them were interconnected by low walls, called parapets. The 3D model is loaded in the AR application in Wavefront OBJ file format and its size is about 5 megabytes.

3.2 Results

The application that has been developed was tested using a Sony VAIO Tap 11 tablet PC with a camera resolution of 1920 x 1080 pixels and yields satisfactory results. The recognition of the pattern object is accomplished regardless of the orientation of the tablet camera and its distance from the pattern object, which ranged from 15 to 25 m, even if only a part of the latter is captured in the camera frame. However, due to the computational burden, the big file size of the 3D model and the high resolution of the video, the frame rate was quite slow. Fig. 2a shows the pattern image, which is an image of a part of the foundations of the south side of the Middle Stoa that remain in site until today. This image with the interest points detected by SURF is illustrated in Fig. 2b; Fig. 2d shows the extracted SURF feature points in a random real world frame, which is shown in Fig. 2c. Fig. 3 indicates the inliers detected by RANSAC, between a random frame and the pattern image, as well as the pattern recognition in a frame. Fig. 4 depicts three random real world frames and their corresponding augmented ones, which constitute the final outcome of the application.

Fig. 2. Pattern image (*a*), detected feature points in the pattern image (*b*), a random real world frame (*c*), and detected interest points in the frame (*d*).

Fig. 3. Inliers detected by RANSAC between a random real world frame and the pattern image (*left*), and recognition of the pattern object in a random real world frame (*right*).

4 Conclusions

An AR application that is intended to improve the overall experience of visitors of the archaeological site of the Ancient Agora of Athens was developed. In spite of the fact that more and more people become familiar with the concept of AR, the application still has an aura of science fiction mystery, which may attract their interest and increase the frequency of their visits to the site. By developing this application, we would like to show the huge power of AR in the visualization of cultural heritage and emphasize the fact that this emerging technology can be used by cultural heritage institutions in order to engage and retain visitors of all ages and educational levels.

The application may be improved in terms of speed, so that the rendering of huge and very detailed 3D models is faster. Furthermore, it may be extended in order to support the import of different 3D models and pattern images and the automatic selection of the appropriate model that will augment the existing situation during visits to various historical and monumental places.

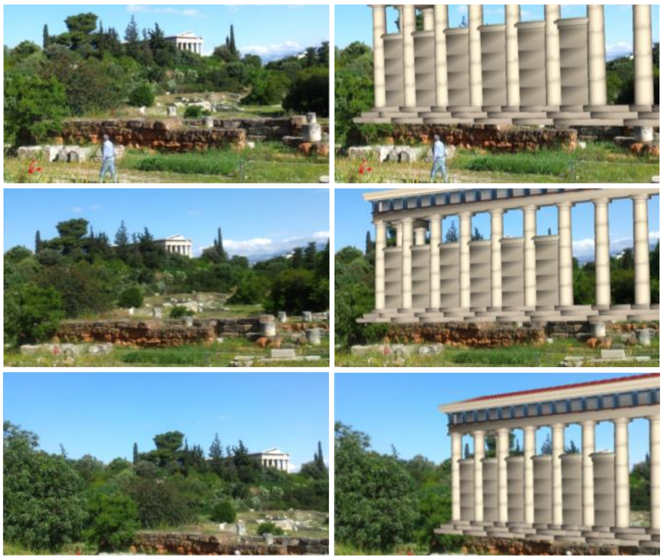

Fig. 4. Real world frames (*left*) and their corresponding augmented frames (*right*).

References

1. Milgram, P., Takemura, H., Utsumi, A., Kishino, F.: Augmented Reality: A class of displays on the reality - virtuality continuum. In: Telemanipulator and Telepresence Technologies Conference of the SPIE International Symposium on Photonics for Industrial Applications, pp. 282—292. SPIE, Washington (1994)
2. Sutherland, I.: A head-mounted three dimensional display. In: AFIPS Fall Joint Computer Conference, pp. 757—764. AFIPS Press, New Jersey (1968)
3. Caudell, T. P., Mizell, D. W.: Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Processes. In: 25th IEEE Hawaii International Conference on Systems Sciences, pp. 659—669. IEEE Press, New York (1992)
4. Verykokou, S., Ioannidis, C.: 3D Visualization Through Planar Pattern Based Augmented Reality. In: 25th International FIG Congress (2014)
5. Azuma, R.: A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environments 6, 355—385 (1997)
6. Kipper, G., Rampolla, J.: Augmented Reality, An Emerging Technologies Guide to AR. Syngress, USA (2013)
7. Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., Ivkovic, M.: Augmented reality technologies, systems and applications. Multimedia Tools and Applications 51, 341—377 (2011)
8. Foni, A. E., Papagiannakis, G. Magnenat-Thalmann, N.: A Taxonomy of Visualization Strategies for Cultural Heritage Applications. ACM Journal on Computing and Cultural Heritage 3, 1—21 (2010)

9. Vlahakis, V., Ioannidis, N., Karigiannis, J., Tsotros, M., Gounaris, M., Stricker, D., Gleue, T., Daehne, P., Almeida, L.: Archeoguide: An Augmented Reality Guide for Archaeological Sites. IEEE Computer Graphics and Applications 22, 52—60 (2002)
10. Papagiannakis, G., Ponder, M., Molet, T., Kshirsagar, S., Cordier, F., Magnenat-Thalmann, N., Thalmann, D.: LIFEPLUS: Revival of life in ancient Pompeii. In: 8th International Conference on Virtual Systems and Multimedia (2002)
11. Granollers, T., Lorés, J., Raimat, G., Junyent, E., Tartera, E.: Vilars. A new dialog model applying Augmented Reality. In: DSV-IS 2002 International Workshop: Design, Specification, and Verification of Interactive Systems (2002)
12. Zöllner, M., Keil, J., Wüst, H., Pletinckx, D.: An Augmented Reality Presentation System for Remote Cultural Heritage Sites. In: 10th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST, pp. 112—116 (2009)
13. EUROMACHS Network for European Digital Media Arts and Cultural Heritage Studies: Augmented Reality, http://euromachs.fl.uc.pt/blog/index.php/tag/augmented-reality
14. Caarls, J., Jonker, P., Kolstee, Y., Rotteveel, J., van Eck, W.: Augmented reality for art, design and cultural heritage: system design and evaluation. EURASIP Journal on Image and Video Processing, 1—17 (2009)
15. Layar: Augmented Reality browsing of Powerhouse Museum around Sydney, http://www.powerhousemuseum.com/layar
16. Museum of London, http://www.museumoflondon.org.uk/Resources/app/you-are-here-app/home.html
17. CultureClic, http://www.cultureclic.fr
18. Netherlands Architecture Institute: Exploring the city with UAR, the NAI's mobile app, http://en.nai.nl/museum/guided_tours/item/_rp_kolom2-1_elementId/1_1023969
19. Boyer, D., Marcus, J.: Implementing Mobile Augmented Reality Applications for Cultural Institutions. In: 15th Annual Museums and the Web Conference (2011)
20. The Connected World: Experiments in Historical Augmented Reality, http://www.lightninglaboratories.com/tcw/2010/11/experiments-in-historical-augmented-reality
21. Haugstvedt, A.C., Krogstie, J.: Mobile augmented reality for cultural heritage: A technology acceptance study. In: 11th IEEE International Symposium on Mixed and Augmented Reality, pp. 247—255. IEEE Press, New York (2012)
22. TAG CLOUD, http://www.tagcloudproject.eu
23. Bay, H., Tuytelaars, T., Van Gool, L.: Speeded-Up Robust Features (SURF). In: 9th European Conference on Computer Vision, pp. 404—417. Springer-Verlag, Berlin – Heidelberg (2006)
24. Fischler, M., Bolles, R.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Communications of the ACM 24, 381—395 (1981)
25. Levenberg, K.: A Method for the Solution of Certain Non-Linear Problems in Least Squares. Quarterly of Applied Mathematics 1, 164—168 (1944)
26. Marquardt, D. W.: An Algorithm for Least-Squares Estimation of Nonlinear Parameters. Journal of the Society for Industrial and Applied Mathematics 11, 431—442 (1963)
27. Zhang, Z.: A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 1330—1334 (2000)
28. Bradski, G., Kaehler, A.: Learning OpenCV. O'Reilly Media, USA (2008)
29. Camera Calibration Toolbox for Matlab, http://www.vision.caltech.edu/bouguetj/calib_doc
30. The Athenian Agora, http://agathe.gr/overview
31. McK. Camp II, J.: The Athenian Agora Site Guide. ASCSA, New Jersey (2010)
32. Agora Monument: Middle Stoa, http://agora.ascsa.net/id/agora/monument/middle%20stoa
33. Travlos, J.: Pictorial dictionary of Ancient Athens. Thames & Hudson, London (1971)
34. ASCSA Digital Collections, http://ascsa.net/research?v=default