Egyptian Journal of Artificial Intelligence
ISSN (Online): 2786-0213, ISSN (Print): 2786-0205
Volume 3, Issue 1, 2024, PP. 13-21

Smart Tour Guide: A Novel Artificial Intelligence System for Replacing Human Guides in Cultural Heritage Sites

Mahmoud H. Abdella[1], Abdelraouf A. Hawash[2*], Ola E. Zahran[3], Ali T. Mohammed[1], Alaa T. Elmaria[4], Yara E. El lakany[1], Mohammed Y. Fathallah[1], Abdelrahman I. Kamal[1], Mona Alnaggar[5], Abdelmawla Yousef[3], Ali I. Siam[3]

[1]Dept. of Data Sciences, Faculty of Artificial Intelligence, Kafrelsheikh University, Egypt

[2]Dept. of Machine Learning and Information Retrieval, Faculty of Artificial Intelligence, Kafrelsheikh University, Egypt

[3]Dept. of Embedded Network Systems Technology, Faculty of Artificial Intelligence, Kafrelsheikh University, Egypt

[4]Egyptian Russian University (ERU), Egypt

[5]Dept. of Robotics and Intelligent Machines, Faculty of Artificial Intelligence, Kafrelsheikh University, Egypt

*Corresponding Author: Abdelraouf A. Hawash, E-mail: abdalraofhawash@gmail.com

Abstract:

Egypt is one of the most popular tourist destinations in the world. In order to enhance the visitor experience and boost tourism revenue, we propose in this paper a smart glasses that serves as a tour guide and a mobile application with remarkable capabilities. The proposed smart glasses are equipped with a voice assistant that allows users to communicate and make inquiries about various topics such as places, hotels, cafes, tourist attractions, and entertainment venues. Additionally, the proposed glasses has the capability to translate Hieroglyphs, natural languages, and even sign language, making them useful for communicating with individuals who are deaf and mute. Furthermore, the glasses can provide navigation and location services, helping users find the shortest routes to their destinations. The device also has the ability to identify historical figures and provide relevant information about them, as well as offer insights into tourist attractions and their histories. Other features include the ability to take pictures, play music, and more. The device can work as a stand-alone system or can be connected to a mobile application for further services. The translation of the hieroglyphic language was accomplished using two YOLOv5 models. The two models achieved a mean average precision (mAP) of 96% and 97.8%, respectively. The tourist sites were recognized using the YOLOv5 Algorithm, with a mAP of 85%. Historical figures, on the other hand, were identified using the YOLOv2 Algorithm, with a mAP of 95%.

Keywords: Smart Tourism; Automatic Hieroglyphs Translation; Historical Figures Identification; Sign Language Translation; Personal Assistant.

1. Introduction

The Artificial Intelligence (AI) approaches have revolutionized innovation in a variety of industries and brought up previously unimaginable possibilities. For instance, using these approaches to classify thyroid diseases [1], identify cardiac conditions [2], recognize facial expressions [3], detect automatic stress [4], detect awareness level [5], detect fatigue state [6] measure heart rate and respiration rate [7].

With modern technologies in the field of AI, especially the field of natural language processing, as well as computer vision, machine learning, and deep learning, its effects at present and in the future will affect many different fields [7].

Egypt is one of the most popular tourist destinations in the world. It is distinguished by its tourism wealth of all kinds, the distribution of temples, museums, monuments, historical and artistic buildings, and vast gardens throughout its territory, as well as its strong infrastructure built to support the tourism industry, which includes hotels, rooms, villages, tourist resorts, tourism companies, and airline offices.

Let us present our solution. It is an easy and convenient way to serve tourists to see, read, and know all the Egyptian antiquities they visit so that they can enjoy an exceptional experience without the need for a tour guide, which increases tourism.

Our solution is a smart tour guide glasses that tourists can use in their trip to ensure having the best experience in Egypt. The smart glasses has several features that tourists need in their trip as follows:

A) Translate Egyptian Hieroglyphs

It is the first and the most important feature of our project, the glasses can recognize texts written in hieroglyphs on the walls of templates and thumbs.

B) Identify pharaohs statues

The second feature in our system is to identify statues of kings and queens of ancient Egypt in tourist attractions. One of the most important ancient Egyptian monuments found in most archaeological sites are the statues of the kings and queens of Egypt, which were carved to perpetuate their achievements. This service helps tourists to identify the king or queen, the owner of the statue and then tourists can search about them and know about pharaohs and their achievements.

C) Landmark and Monuments recognition

The third feature in our system is Landmark Detection. This feature enables tourists to identify the monuments and thus they can learn more information about these monuments and their history by searching with the voice assistant.

D) Voice assistant and translating natural languages

This is one of the most exciting features in our system. Voice assistant helps tourists in their trips and facilitates communication between the tourist and the glasses. It helps users to use the glasses via voice-commands and get response as speech instead of text.

It can help the tourist in many things such as: telling the time and the weather, searching about something on Google or Wikipedia, opening videos on YouTube, playing music, and translating languages to help people understand hieroglyphics.

E) Sign Language detection

The last feature in our system is sign language detection. This feature has been added to help people speak with sign language to communicate with others easily. It helps tourists to have a better experience. This can be very helpful for the deaf and dumb people in communicating with others as knowing sign language is not something that is common to all.

2. Related Studies

modern technologies have aided the preservation, analysis, and study of cultural, historical, and artistic content. For example, wavelet analysis of brush strokes in paintings can reveal the artist for that painting [22]. Furthermore, the digital Michelangelo project [23] created high quality 3D models of Michelangelo's sculptures and architecture. In this project I presented a new dataset, and visual approach for automatic Egyptian hieroglyphs detection and recognition and use the results of this in the work of a translation of the text written in hieroglyphs.

Related work on automatic hieroglyph recognition focuses on Mesoamerican culture, and on the ancient Maya hieroglyphs [24, 25, 26], there is also work related to translating hieroglyphs from images and returning them as text [1]. To this end, I suggested using YoloV5 algorithm to identify and recognize the symbols, and then presented an algorithm to arrange

the hieroglyphic symbols in a way that achieves the correct meaning as shown in figure 1.

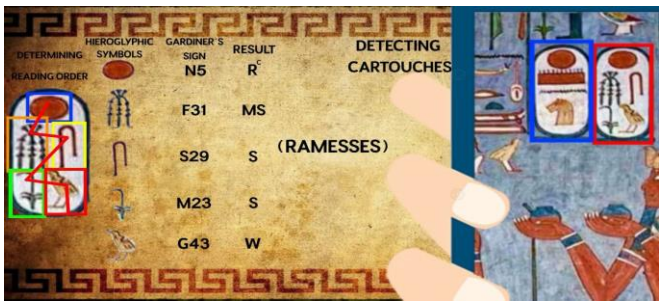

Figure 1: The sequence of work of the hieroglyphic translation algorithm.

Related work on recognizing an ancient statue faces for 3 kings using haar-cascade and ANN [14,15,16]. In this paper we propose a new approach to recognize pharaoh statues faces in real time for more than 25 pharaohs using CNN and Face Recognition module as shown in figure 2.

Figure 2: Recognizing historical figures.

3. Proposed System

In this paper our goal is to design and develop a Smart Tour Guide device for tourists visiting ancient Egyptian sites. The Smart glasses tour guide is a wearable device that will be equipped with advanced hardware and software technologies that enable it to recognize and translate ancient Egyptian language and provide instant translations. The system will also implement face recognition technology to provide accurate information about the ancient Egyptian figures. and offers many other services. The system is also able to identify historical monuments. The system is also supported with voice assistant to help tourists on their trip. The system also supports the recognition of sign language to improve the experience of its speakers and facilitate communication with them.

3.1 Materials

This project contains some component such as:

- **Sipeed Maixduino Kit RISC-V**:

SIPEED MaixDuino is a development board compatible with Arduino based on MaixPy M1 module (main control: Kendryte K210). It integrates camera, TF card slot, user buttons, TFT display, MaixDuino expansion interface, etc. look at figure 3. users can use MaixDuino to easily build a face recognition access control system, and reserve development and debugging interfaces, which can also be used as a powerful AI learning development board [10,19,20,21].

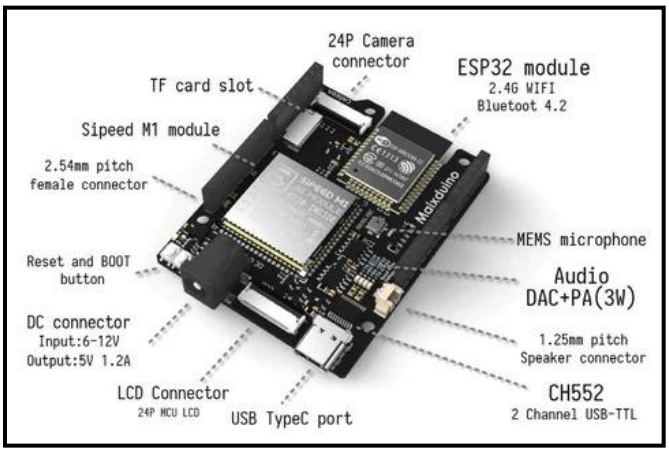

Figure 3: Sipeed Maixduino Kit RISC-V .

- **Display:** A small display that can be placed in front of the user's eyes to provide visual information.
- **Camera:** A camera can capture visual information about the surroundings and enable features such as object recognition and augmented reality.
- **Audio components**: The glasses can include speakers and microphones to provide audio information and enable voice commands.
- **Wi-Fi and Bluetooth:** Wi-Fi and Bluetooth can help the glasses to connect to the internet and other devices, such as smartphones or smartwatches.
- **Battery**: A rechargeable battery can power the device for an extended period.

All components used in the system are shown in figure 4 .

Figure 4: Smart tour guide component.

3.2 System implementation

In this project we implement some important features:

3.2.1 translate Egyptian hieroglyphs

In this system, for this feature we propose a powerful strategy for translating hieroglyphs with two main steps:

Step 1: Detect cartouche in the hieroglyphic mural:

- First the model detects all cartouches in the mural and calculates the area for each cartouche as shown in figure 5.
- Converts each cartouche to a separate image. The model saves these images to be the input for second module. It is an object detection task. We used YOLOv5 to detect cartouche in the hieroglyphic mural [11].

Figure 5: detect cartouches in mural.

Step 2: Detect and Translate names in cartouche:

second step in to detect and recognize signs in the Cartouche and arrange them to construct the name or the word as shown in figure 6.

The steps of translating name as flow:

1. detecting the signs in the cartouche using YOLOv5.

Figure 6: detect signs in cartouche.

2. then arrange signs to construct the name correctly:
   - First the model starts to arrange from top to bottom Based on Y coordinates as the symbol with the minimum vertical coordinate value comes first.
   - If there are two signs that intersect vertically, it looks for x-coordinates for two signs. A check is performed on their horizontal coordinates. If the minimum value of the horizontal coordinate of one symbol is greater than or equal to the maximum value of the horizontal coordinate of the other symbol, the symbol with the larger horizontal coordinate takes precedence and is placed before the other symbol.

- The process continues until all symbols are sorted based on these criteria as shown in figure 7.

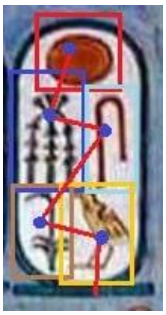

Figure 7: arrange signs in cartouche.

3. Finally, the model output the name in cartouche in English.

3.2.2 Identify pharaohs statues

In this system, we propose to use a Deep Model for face detection and recognition to recognize the faces of pharaohs in the history background [13,18].

- First, we collected data for pharaohs statues, used data augmentation to increase the data and ensure to collect the faces of the king's statue in different poses.
- For detect faces in statues, we used pretrained CNN model that perform more accurate than other techniques.
- We changed each face in the training dataset to an encoded array which contains facial landmarks.
- In the last stage the face of the statue will be compared to the well-known encoded facial training datasets.
- The closest match is found by using the Euclidean distance which finds the minimum difference between the weights of the input image and the set of weights of all images in the database as depicted in figure 8.

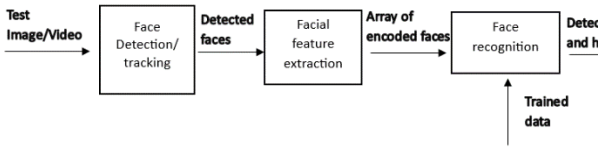

Figure 8: steps of the face recognition model .

- After recognizing the king's name, it is optional to show information about him from database or search about him on Google using Voice assistant.

3.2.3 Landmark and Monuments recognition

Monument recognition is a challenging problem in the domain of image classification due to huge variations in the architecture of different monuments. Different orientations of the structure play an important role in the recognition of the monuments in their images [4,5].

As depicted in figure 9, we propose an approach for classification of various monuments based on the features of the monument images as flow:

- For this system, we gathered images for 10 different Egyptian Monuments. As lighting conditions and different angles of a monument are two factors that directly affect the coloring and possibly the shape of a monument, we try to maintain a variety in our images. With this strategy, it is ensured a greater chance for a correct and accurate prediction, even under various circumstances.
- It is important to ignore very low-resolution images as the information gained from them is minor and consequently, the overall performance is decreasing.
- For images annotation, we used a tool named LabelImg enabling us to easily draw of the desired bounding box along with the annotation for each box. LabelImg also offers us the opportunity for saving the annotation with either YOLO or PascalVOC format. And we can make annotation with Roboflow.
- Then, we use YOLO algorithm to build our model. We used YOLOv5. We trained our model on the 10 classes in training data.
- The model can detect the monument in the image and recognize it with high accuracy and high confidence.

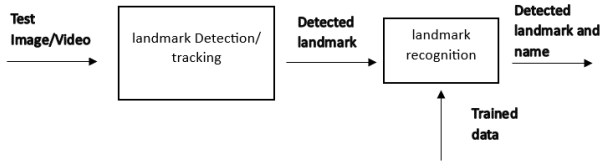

Figure 9: steps of the landmark detection model.

3.2.4 Voice assistant and translating natural languages

Voice assistant in this system works as depicted in figure 10:

1. First it can listen to human voice, save it and convert it to text using Speech recognition Module.
2. Based on text generated, voice assistant defines which function to do and do it.
3. The output generated as text. Voice assistant converts it to speech again using TTS Module and out it to human.

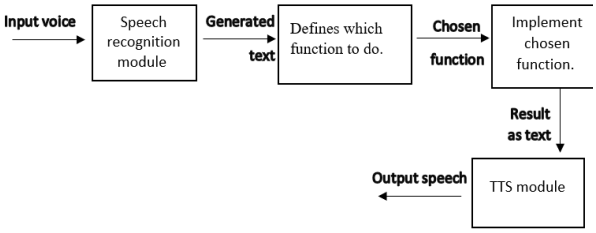

Figure 10: steps of full Voice assistant .

Voice assistant can do multiple functions. The most important usage for Voice assistant is its ability to translate natural languages [2,3,8,17], as depicted in figure 11.

How voice assistant work:

1. First, the Voice assistant service should be open. Then the user should call the translating function.
2. For translation, the first user chooses the language to which translate.
3. The service uses a translator function from Googletrans module to translate the input text.
4. Using Google-Text-to-Speech i.e., gTTS() method to speak the translated text into the destination language.

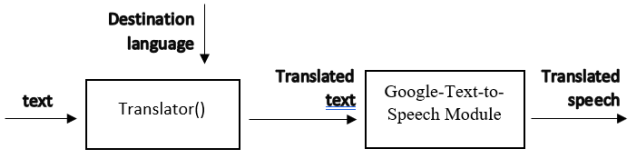

Figure 11: steps of translation.

3.2.5 Sign Language detection

In this system, we propose to use a Deep Model for detecting and recognizing hand sign language [9].

- First, we collect 300 images for each letter, all of which are hand shaped in different positions.
- We used HandTrackingModule to detect hands in real-time as depicted in figure 12.
- The MediaPipe Hand Landmarker task detects the landmarks of the hands in an image. It helps to localize key points of the hands and render visual effects over the hands.
- We used ClassificationModule from CVZone with CNN model to classify the hand detected and recognize the letter.
- The model works on the bound boxes for hands detected in hand detector module, and for bound box it predicted the most match class for it.
- Our model had been trained to successfully recognize and classify 15 sign language characters, and that is very good in this case because Sign language is very difficult, and its signs are somewhat similar.

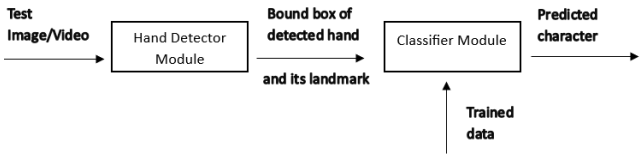

Figure 12: steps of sign language detection.

4. Results

We introduced a smart tour guide who can enrich the tourist experience with the help of artificial intelligence.

Not only do smart tour guides offer real-time navigation and translation capabilities, but they can also provide historical and cultural context to landmarks and attractions. This can greatly enhance the overall tourism experience by offering a deeper understanding of the places being visited.

• Our Smart glasses has some important features, including the personal assistant who takes the question from the user as a voice, converts the speech into text, performs a search on Google, and returns to the user the required answer as text.

• The second feature is the recognition of the face, as it identifies the faces of kings and pharaohs, and then gives them information about this king.

• The third feature is QR Code, as it can read the QR code or barcode in a simple way.

• The fourth feature is capturing photos, as it takes a souvenir photo of the tourist in high quality and sends this photo to the server so that you can print this photo.

• The fifth feature, which is the most important feature of our project, is that it can recognize texts written in hieroglyphs.

• And in anticipation of communicating with one of the disabled who is deaf and dumb, the sixth feature comes as the ability to recognize sign language.

• Last feature is Landmark detection is the process of finding significant landmarks in an image. We used it in our project in order to identify the archaeological monuments in the museum and talk about their names and some information about them.

Overall, while there are potential benefits and challenges associated with using smart glasses as a tour guide, they offer a unique and exciting way to enhance the tourism experience for both tourists.

5. Conclusion

We introduced smart glasses and a mobile application that acts as a tour guide with exceptional features to contribute to improving the tourist experience, which increases tourism income[6]. This system is equipped with a voice assistant that the tourist can talk to and inquire about what he wants from places, hotels, cafes, tourist attractions and entertainment places, not to mention Its ability to translate the hieroglyphic language to be able to understand the ancient Egyptian texts, as well as the translation of natural languages and even sign language in anticipation of dealing with a deaf and dumb person, in addition to the ability of the system to determine the places and the shortest paths that lead to your destination, as well as the possibility of identifying historical figures and providing the tourist with the necessary information We will work in the future to develop the project so that it can translate the entire hieroglyphic text and identify all historical figures and all tourist attractions, in addition to generating actual images of historical figures using deep learning algorithms.

Competing interests: The authors declare that they have no competing interests.

Authors contributions: The authors declare that they contributed equally to all sections of the paper. All authors read and approved the final manuscript.

References:

1. Alnaggar M, Handosa M, Medhat T, Z Rashad M. Thyroid Disease Multi-class Classification based on Optimized Gradient Boosting Model. Egyptian Journal of Artificial Intelligence. 2023 Apr 1;2(1):1-4.

2. Alnaggar M, Handosa M, Medhat T, Rashad MZ. An IoT-based Framework for Detecting Heart Conditions using Machine Learning. International Journal of Advanced Computer Science and Applications. 2023;14(4).

3. Alhussan AA, Talaat FM, El-kenawy ES, Abdelhamid AA, Ibrahim A, Khafaga DS, Alnaggar M. Facial Expression Recognition Model Depending on Optimized Support Vector Machine. Computers, Materials & Continua. 2023 Jul 1;76(1).

4. Siam AI, Gamel SA, Talaat FM. Automatic stress detection in car drivers based on non-invasive physiological signals using machine learning techniques. Neural Computing and Applications. 2023 Jun;35(17):12891-904.

5. El-Rashidy N, Sedik A, Siam AI, Ali ZH. An efficient edge/cloud medical system for rapid

detection of level of consciousness in emergency medicine based on explainable machine learning models. Neural Computing and Applications. 2023 May;35(14):10695-716.

6. Alharbey R, Dessouky MM, Sedik A, Siam AI, Elaskily MA. Fatigue state detection for tired persons in presence of driving periods. IEEE Access. 2022 Jun 22;10:79403-18.

7. Alnaggar M, Siam AI, Handosa M, Medhat T, Rashad MZ. Video-based real-time monitoring for heart rate and respiration rate. Expert Systems with Applications. 2023 Sep 1;225:120135.

8. Morris Franken and Jan C. van Gemert, "Automatic Egyptian Hieroglyph Recognition by Retrieving Images as Texts", in 2013, Intelligent Systems Lab Amsterdam (ISLA), University of Amsterdam Science Park 904 Amsterdam, The Netherlands.

9. Basem H. A. Ahmed, Ayman S. Ghabayen," Arabic Automatic Speech Recognition Enhancement", in 2017, Palestinian International Conference on Information and Communication Technology

10. Ankit Pandey, Vaibhav Vashist, Prateek Tiwari , Sunil Sikka, Priyanka Makkar, "Smart Voice Based Virtual Personal Assistants with Artificial Intelligence", in 2020 , Amity School of Engineering and Technology, Haryana, India.

11. Aradhya Saini, Tanu Gupta, Rajat Kumar, Akshay Kumar Gupta,Monika Panwar, Ankush Mittal, "Image based Indian Monument Recognition using Convoluted Neural", in 2017, 2017 International Conference on Big Data, IoT and Data Science (BID).

12. Aman Preet Gulati, "Hand landmarks detection on an image using Mediapipe", in 2022, Analytics Vidhya

13. Ahmed K Elnagar , Abdelkader Mohammed Sghaier Derbali , "The Importance of tourism contributions in Egyptian economy", in September 2020, https://www.researchgate.net/publication/344013962_The_importance_of_tourism_contributions_in_Egyptian_economy

14. Simplilearn," Has Technology Improved Our Lives?", in 2023,https://www.simplilearn.com/how-has-technology-improved-our-lives-article .

15. What is Natural Language Processing? | IBM. (n.d.). https://www.ibm.com/topics/natural-language-processing.

16. IBM, "What is Computer vision?", What is Computer Vision? | IBM. (n.d.). https://www.ibm.com/topics/computer-vision.

17. Sipeed Wiki, "Introduction of MaixPy", in 2018, https://maixduino.sipeed.com/en/get_started/get_hardware.html.

18. Rohit Kundu,"YOLO: Algorithm for Object Detection Explained [+Examples]", in 2023, https://www.v7labs.com/blog/yolo-object-detection

19. Zoumana Keita,"YOLO Object Detection Explained", in 2022, https://www.datacamp.com/blog/yolo-object-detection-explained

20. kaspersky, "What is Facial Recognition – Definition and Explanation", https://www.kaspersky.com/resource-center/definitions/what-is-facial-recognition

21. Ahmed Hisham Misbah, "Can computer Vision recognize an ancient statue face?", in 2019, Research department at Artificial intelligence technology centerMisr university for Science and technology

22. Ali Tarhini, "Face Recognition: An Introduction", in 2010, https://alitarhini.wordpress.com/2010/12/05/face-recognition-an-introduction/

23. Firas Husham Al-Mukhtar,Mustafa Zuhaer Nayef Al-Dabagh, "Real-Time Face Recognition System Using KPCA, LBP and Support Vector Machine" in 2017 ,International Journal of Advanced Engineering Research and Science (IJAERS)

24. Benjamin Pryke,"How to Use Jupyter Notebook: A Beginner's Tutorial", in 2020, dataquest blog, https://www.dataquest.io/blog/jupyter-notebook-tutorial/

25. https://pypi.org/project/face-recognition/

26. MaixHub, "AI online training and sharing platform" https://maixhub.com/home

27. Maixduino example sketch "selfie" not working, in 2020, https://forum.arduino.cc/t/maixduino-example-sketch-selfie-not-working/667685

28. MaixPy_scripts on GitHub, in 2019, https://github.com/sipeed/MaixPy_scripts

29. R. Johnson, E. Hendriks, J. Berezhnoy, E. Brevdo, S. Hughes, I. Daubechies, J. Li, E. Postma,

and J. Wang. Image processing for artist identification. Signal Processing Magazine, IEEE, 25(4):37–48, 2008.

30. M. Levoy et al. The digital Michelangelo project: 3D scanning of large statues. In Computer graphics and interactive techniques, 2000.

31. Y. Frauela, O. Quesadab, and E. Bribiescaa. Detection of a polymorphic mesoamerican symbol using a rule-based approach. Pattern Recognition, 39, 2006.

32. E. Roman-Rangel, C. Pallan, J.M. Odobez, and D. Gatica-Perez. Retrieving ancient maya glyphs with shape context. In ICCV Workshop, 2009.

33. E. Roman-Rangel, C. Pallan Gayol, J.M. Odobez, and D. Gatica-Perez. Searching the past: an improved shape descriptor to retrieve maya hieroglyphs. In ACM MM, 2011.