

Visualization Learning for Visually Impaired People

Patrick Salamin, Daniel Thalmann, and Frédéric Vexo

VRLab at EPFL, Switzerland,
WWW home page: <http://vrlab.epfl.ch/>

Abstract. Virtual Environments (VE) are mainly visual experiments that exclude visually impaired people. In this paper we present an application that should allow almost everybody to “see” or at least to perceive 3D shapes. We will first describe the mandatory aspects of such an application, the tests we made and finally conclude by the results we obtained with geometrically basic shapes that are very promising.

1 Introduction

Access to digital content, e.g. the World Wide Web, is mainly based on the visual modality. Since 1997, through the WAI¹ program, the W3C aims to render the Web content accessible by everyone: this means that visually impaired people do not have to be excluded. Authors in [16][7] presented some works which aim to solve these accessibility issues. Therefore P. Roth et al. showed in [11] that several problems still remain. As stated by Ardito et al. in [3], this accessibility aspect is especially crucial for e-learning applications.

In this paper, we will focus on the access to digital 3D shapes by visually impaired people (VIP) in the learning context. Because of the importance of the visual / geometric reasoning in the learning procedure[9][1], it is quite difficult to convey this type of information to the VIP and they have not to be born blind because they need to recognize basic 2D geometrical shapes. Audio haptics and touch efficiency have already been applied in [6] for similar situations with blind people with promising results. This is why we decided here to experiment the trans-modality mechanisms. In order to validate this concept, we have developed an application giving the possibility to visually impaired people to understand 3D shapes stored in digital libraries on the Web with the help of audio haptic feedback. To ensure a large dissemination, our system has to remain low-cost[2]. For that, we have developed an application based on hand-held devices (e.g. a PDA which is common to use and quite cheap) with Internet connection. This will allow users to perceive the contour of the 2D views of the 3D shapes and their volume by tonality variations.

The first section will present a brief overview of the current projects and research targetting the accessibility and the visualization based on audio, tactile

¹ <http://www.w3.org/WAI>

and force-feedback haptics for the content of digital libraries. The second one will present then our transmodal rendering tool allowing VIP to understand digital 3D contents. We will conclude with a description of the tests we performed to fairly evaluate our application. Finally the results we obtained and the possible improvements will be discussed.

2 Related work

The perception of 3D shapes is crucial in Education [4]. Didactic material becomes more and more digitalized. Authors in [11] show that the access to digital content is difficult or even not suitable for disabled or visually impaired people. Therefore, several research works have been performed to solve this issue with the help of haptic feedback created on the remaining available senses. These works could be categorized into three groups: haptic force feedback, tactile and audio haptics.

If the visual modality is not available, the authors in [13] affirm that the haptic force feedback is the best to perceive 3D shapes, except for the exploration of sharp edges and corners because of the loss of contact with the shape. As presented in [10], it would be useful to couple force feedback with another modality in order to obtain acceptable results. We can then conclude that haptic force feedback allows to perceive only a certain kind of 3D shapes with rounded, smooth edges. However, force feedback devices are invasive, expensive and restrain the workspace to a very limited area and this is in contradiction with our requirements for large public Education.

Since haptic force feedback remains too expensive, invasive and would prevent from a large dissemination, some authors explored other aspects of the sense of touch: tactile haptics, which is powerful to render fine touch. Concerning the research made in this domain, we could mention tactile displays which, according to the authors in [18], allow a good size differentiation of surface shapes (e.g. the NIST² Braille Reader). These particular screens also seem to be a very good solution for web surfing or mail checking while the size of the page to render is small enough [14]. We assume then that the surfaces can easily be covered by the user with tactile haptics, but it seems difficult to represent a volume or even to give enough depth information on the shape. Based on this, we can conclude that haptic tactile seems to be perfect to cover a surface and give texture information, but it does not provide volume information. Moreover, due to the non-portability of haptic tactile devices – because of their weight, their need to be linked to a computer and the impossibility to make them work efficiently with batteries –, a large dissemination of our system would be compromised.

Due to its workspace limitations, its invasiveness and its price, haptic force feedback does not seem to correspond to our requirements: a mobile, low-cost and efficient system allowing a large dissemination. Tactile haptics would be interesting for surface covering but it needs a particular hardware which is not

² <http://www.nist.gov>

easily portable. Researchers have then looked for another modality: haptic audio. Concerning the perception of the environment, we can mention obstacle avoidance applications [5] and 3D audio systems [8] built for a better geographical location of the subject in the world. Furthermore, the authors in [12] gave the possibility to move and build in the shapes for a better environment recognition. The results obtained in this last research would be easily applicable to represent a virtual world and then recognition of virtual shapes within the framework of visually impaired people Education. Nevertheless the main problem with the audio haptics resides in the user interaction with the environment, e.g. to know where obstacles are. If we consider that the user hands indicate the position and direction the user wants to point to their location and orientation depending on the virtual world needs to be known, which is very difficult to obtain with enough accuracy and needs a strong user training. We can then conclude that the haptic audio is powerful to render volume information, but its main problem resides on the surface covering which is almost impossible.

According to these previous works, to combine modalities[17] and especially audio feedback and tactile feedback seems the most promising[15]. Moreover, as we want a low-cost system for a large dissemination, the smartphone or the PDA seems the best device to develop our system because it provides a tactile screen to cover the shape and loudspeakers to render the volume information. The next section will present the tool we have developed to allow VIP to perceive digital 3D shapes.

3 Tools for 3D shapes description

According to the results shown in the previous section, we present hereinafter our multi-modal application that allows visually impaired people to discover and represent themselves 3D shapes stored on the WEB. They only need a hand-held device with a Wi-Fi connection (Fig. 1 on the left). We will first present the data extraction and preparation from the 3D shape and then their rendering for the user shape perception, the tests we made and their results.

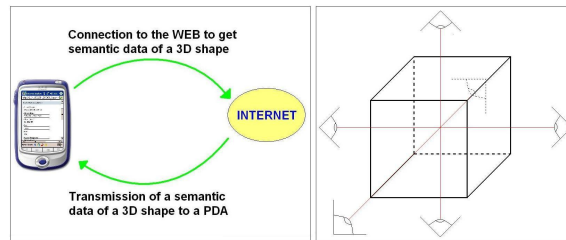


Fig. 1. On the right: Easy acquisition of the data; on the left: Camera settings for the snapshots

3.1 Semantic extraction of the data

Semantic extraction of useful data from the 3D shape to allow a multi-modal rendering for visually impaired people is performed on the server before any user connection (PDA or smartphone). For this, we first load the shape in a simple 3D OpenGL-based viewer in order to take snapshots and other interesting information like the associated depth buffer.

Concerning the points of view (or camera settings) for the snapshots we currently take six snapshots every 90 degrees (corresponding to the six faces of a bounding box), as you can see on the right of the Figure 1. The user performs a mental reconstruction after he/she has explored every six faces in order to re-build the volume.

Notice that as VIP are not necessarily blind, we apply some operations on the snapshots with the help of the associated depth buffer to increase the contrast. The obtained picture we present hereinafter could help these non-blind VIP to approximate the shape. Concerning the contrast, in order to detect if a pixel is in the contour of the shape or in the form, we check its corresponding value in the associated depth buffer. If the value shows that the pixel is outside of the shape, it becomes black to indicate it as part of the background. Otherwise, it is considered as a pixel in the shape and becomes then white. We finally provide six pictures of a white shape face on a black background to the VIP.

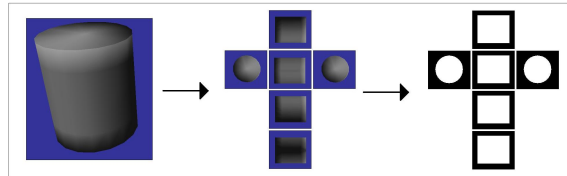


Fig. 2. Semantic data extraction

Once the pictures and associated depth buffers are stored, any user can get them via the Internet and work in local after the download.

3.2 Semantic rendering of the data

Once the extracted data are downloaded on the smart phone, the user begins to scan the tactile display with the PDA stylus to recognize the shape face displayed on the full screen. In this application, due to the targetted public, the graphical interface is almost inexistent. This sobriety allows the user to concentrate on the shape recognition process that we describe hereinafter.

In a first step, the user understands the contour of the current face while scanning the tactile display. A song tonality is indeed produced when the stylus is pointing on the shape. For this, we check the associated depth buffer value at the stylus position (Fig. 3). In a second step, depending on the shape volume

– i.e. the depth buffer values – the users will remark that the sound tonalities change. They can then understand the convexity of the face. Once the contour and the curvature of the face are completely discovered, the users can go to another face with the help of the joystick buttons of the smartphone.

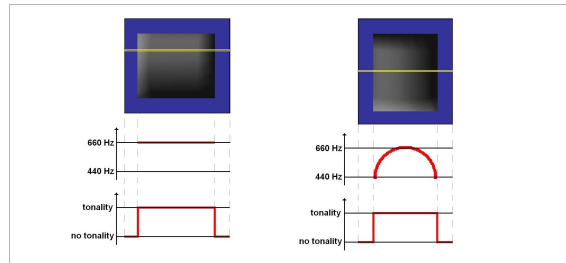


Fig. 3. Continuous tonality variation depending on the shape

After having explored every face, the user should be able to recognize the current 3D shape with limited memory efforts. We will present in the next section the tests we have performed to prove the efficiency of our application and the results we obtained.

4 Tests to validate the application

In order to validate our previously described application, we have performed some experiments described hereinafter. 3D shapes are provided to the testers who use different scanning method of the display. Our efficiency criteria are based on the shape recognition, the elapsed time and the user feeling. We have made the assumption that every shape should be relatively quickly recognized by the testers.

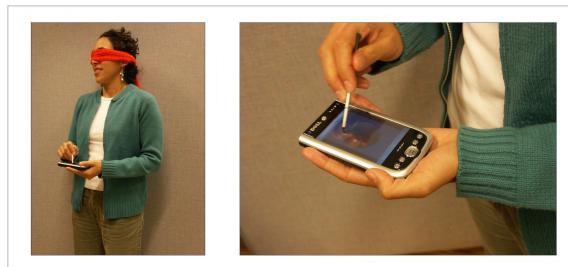


Fig. 4. User testing the application

4.1 Presentation of the tests

To perform these experiments, we worked with ten testers from twenty to forty years old. They are seven males and three females. We simulate the worst case to validate our tests: we worked with non-visually impaired people with a blind-fold. This means that, as the blind people, they cannot see the PDA (Fig. 4). Moreover, there is an added difficulty to perform the shape recognition because they are not used to work with only audio feedback. At each step, the forms is randomly presented to the testers.

We tested a panel of 3D shapes which are basic geometrical shapes containing angles, curves and plane surfaces: a sphere, a cylinder, a cube, a pyramid, a cone, a tetrahedron and a hourglass (two cones joined at their top). This panel (see Figure 5), composed of several forms offering very different characteristics, seems to be representative enough to validate our system for the geometrically-basic 3D shape.

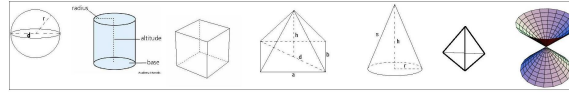


Fig. 5. Panel of the seven proposed shapes for the experiments

Once the shape are downloaded on the smartphone, the users begin to discover it by scanning the display with the stylus. According to the experiments and some observations we made up to now, we have discovered that there seem to be three main way to scan a display to get a form. The first solution, consisting in scanning the display in a random way, is usually the less efficient. Otherwise, the users sometimes try to find where is a part of the shape (or the shape) and trace horizontal (or vertical) lines while they are on the shape and discover it very quickly. But if there are two shapes on the screen, they would miss the second one. The third method is very similar to the second except that the testers scan the full screen to discover everything on the display. After some tests, we assume that this last one would be the most efficient.

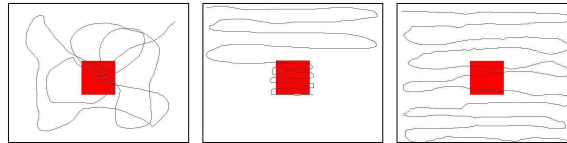


Fig. 6. More common way to discover a shapes by navigating on a screen

Finally, the objective criteria we use to validate the experiments for our application are first based on the result of the experimentation (does the subject

recognized the shape or not), the elapsed time while he/she was guessing it and his/her feeling (how hard he/she found this experiment, intuitive or not, any remarks). Obviously, the order of appearance of the shapes during the test are randomly chosen. We will then present the results we obtained during these experiments in the following section.

4.2 Results

Different results obtained for each shape are presented hereinafter. We have three criteria of validation: two conditional depending on each other for the usability of the application and a time-dependent one for its efficiency. If the tester recognize the shape, the first conditional criterium is full-filled. Otherwise, we propose to him/her a panel of shapes and the second conditional criterium depends on his ability to find the shape with this help. The last criterium concerns the time needed by the user to recognize the shape.

As written in the previous section, during the tests, the shapes appears in a random order. This explains why some more experimented users find faster the name of the shape comparing to others whose the current shape was one of the first to discover. We will now describe the experiments for each shape with a diagram showing what the tester listened while he scanned the tactile display as shown with the red segment for the most representative of them.

In the case of the cube, the first conditional criterium was always reached because every tester found the shape. Concerning the efficiency criterium, the average time taken by the users to recognize the cube is around one minute which means that for this 3D shape, our method is quite efficient. Notice that none of the testers had any hesitation for the recognition of the cube.

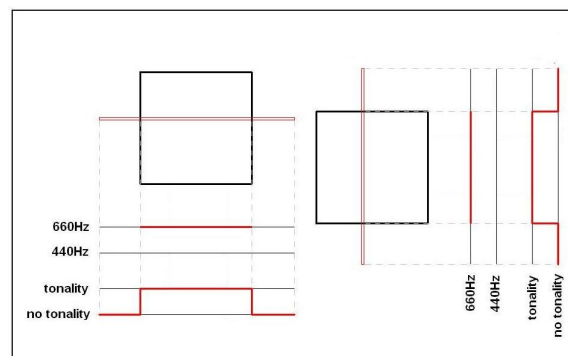


Fig. 7. Representation of the schemas and associated tonalities obtained during the experiment of the cube

As we foresaw it, they first discovered a plane square face, a second one and most of them already said “It is a cube!”. As you can see on the Fig. 7, there

is no tonality while the stylus is outside of the shape – a square. Moreover, as the cube face is parallel to the display, the tester listens to a monotone tonality (no variation) while he/she scans the shape. On the following Figures (8, 9 and 10), we present a diagram of the provided tonalities for some proposed shapes in order to clearly explain our method. Concerning the recognition of the sphere, every tester full-filled the usability criterium by finding the good shape. The time needed by the testers varied between forty-five and ninety seconds. For the case of the regular cylinder, every tester recognized the shape between fifty and eighty seconds without any hesitation. The square-based pyramid was a troubling shape for half of the users because of its plane surfaces that are not “parallel” to the PDA screen. This means that tonality varies only while scanning vertically (instead of horizontally). Finally every user recognized the shape, spending around ninety seconds.

The cone is a pretty complicated shape to recognize when we are not able to see it because with the horizontal scanning, we can understand the convexity of the face (see Fig. 8), but we also have a constant tonality variation when we scan vertically the screen which seems to introduce lots of doubts into the testers mind. The first criterium of usability is finally full-filled by each tester for this shape, but the efficiency one is not very conclusive: an average time of one hundred and fifty seconds is needed by the users to recognize the shape.

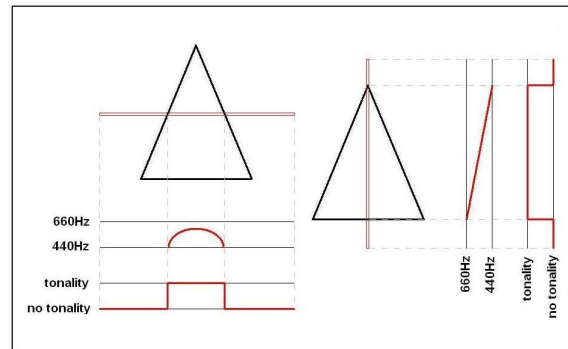


Fig. 8. Representation of the schemas and associated tonalities obtained during the experiment of the cone

The tetrahedron is in our sense one of the most confusing shapes proposed in our tests because of no parallel face to the PDA screen (see Fig. 9). Seven of the ten people guess and find it after three hundred seconds and two others deduct it directly with the help of the panel of shapes. The depth information seemed to be confusing. The conditional criteria are then not full-filled for every testers and the time taken by the testers is quite long for a “basic”, but not so common shape.

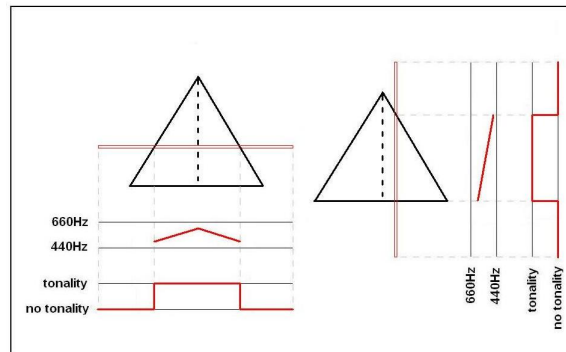


Fig. 9. Representation of the schemas and associated tonalities obtained during the experiment of the tetrahedron

For the case of the hourglass, eight testers recognize it in less than one hundred seconds because of its contour (see Fig. 10). Most of the testers recognized it without any help in less than one hundred seconds.

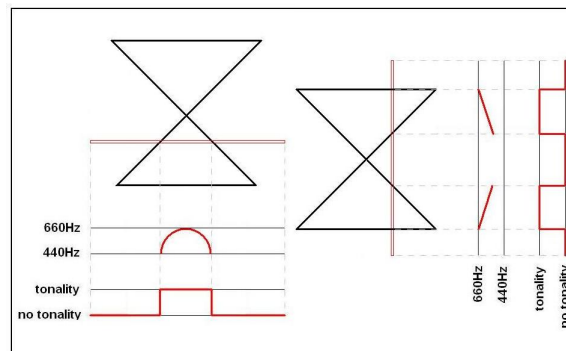


Fig. 10. Representation of the schemas and associated tonalities obtained during the experiment of the hourglass

Finally, we compare the average times needed by the testers to recognize each shape (Fig. 11) and conclude. We can notice that each shape was recognized by almost every tester even if they are all non-blind people. This could explain why they needed so much time to discover the shapes. Moreover, so far as we foresaw that the tetrahedron would be the hardest-to-recognize shape, on the reverse the hourglass was very easily and quickly understood by every tester. Our system seems thus to be quite efficient.

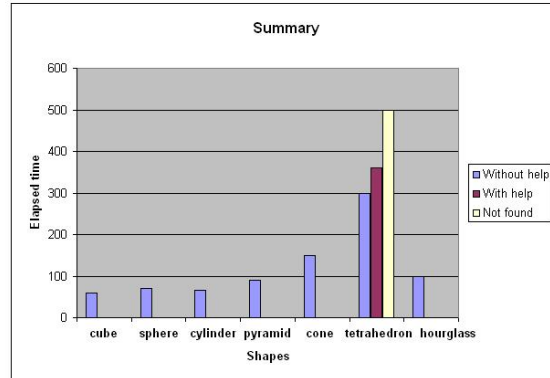


Fig. 11. Table summarizing the obtained results of the previous experiments

5 Conclusion

We can conclude that our application seems to be efficient. Based on the tests and results obtained in this paper, we can conclude that we have built an efficient application allowing people to understand geometrically basic 3D shapes without or with a very reduced vision.

With the help of our application, the users can first easily understand the contour of the pictures, corresponding to predefined views we chose and propose them. Second of all, the concept of continuous tonality variation depending on the height of the shape seems to be very intuitive and easy to use. Finally, the last step, which consists in mentally assembling all the faces to re-build the complete 3D shape, does not seem to be difficult at all for any tester. We can then conclude our system seems to be an efficient and promising way for visually impaired people to be able to understand 3D shapes. We will propose hereinafter future improvements and interesting tests we would make.

Nevertheless, future works with more complex shapes (e.g. a person face, a car or any common object) would be very interesting. It would then be very useful to have a digital library on the Internet available for everyone containing 3D models and their 2D views we could load on our hand-held device.

Concerning the provided views of the 3D shape, we plan some collaboration to work with a “best view” algorithm. It will provide the best positions and angles for the cameras in order to get the most efficient snapshots for our application. Notice that in this case, we should find a way to advise the user the camera settings for the snapshot (instead of our views where these information are obvious).

We plan as final improvement to make all these experiments with visually impaired people in order to highlight the problem they would encounter with our system and finally certify the efficiency of our application.

6 acknowledgments

This research has been partially funded by the Swiss Federal Office for Education and Science under the framework of the European Network of Excellence IST-AIM@SHAPE³.

References

1. A.M.Barry, D. Berry, S. Cunningham, J. Newton, M. Schweepe, A. Spalter, W. Whiteley, and R. Williams. Visual learning for science and engineering. In *Visual Learning Campfire Symposia Program, ACM SIGGRAPH and Eurographics co-sponsored*, June 2002.
2. G. B. Anderson and D. W. Rogers. An inexpensive braille terminal device. *Commun. ACM*, 11(6):417–418, 1968.
3. C. Ardito, M. D. Marsico, R. Lanzilotti, S. Levialdi, T. Roselli, V. Rossano, and M. Tersigni. Usability of e-learning tools. In *AVI '04: Proceedings of the working conference on Advanced visual interfaces*, pages 80–84, New York, NY, USA, 2004. ACM Press.
4. N. Baloian, W. Luther, and J. Sanchez. Modeling educational software for people with disabilities: theory and practice. In *Assets '02: Proceedings of the fifth international ACM conference on Assistive technologies*, pages 111–118, New York, NY, USA, 2002. ACM Press.
5. S. Cardin, D. Thalmann, and F. Vexo. Wearable system for mobility improvement of visually impaired people. In *Visual Computer journal*, 2006.
6. A. Chang and C. O'Sullivan. Audio-haptic feedback in mobile phones. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1264–1267, New York, NY, USA, 2005. ACM Press.
7. M. Y. Ivory, S. Yu, and K. Gronemyer. Search result exploration: a preliminary study of blind and sighted users' decision making and performance. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1453–1456, New York, NY, USA, 2004. ACM Press.
8. M. Lumbreras and J. Sanchez. Interactive 3d sound hyperstories for blind children. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 318–325, New York, NY, USA, 1999. ACM Press.
9. M. McGrath and J.R.Brown. Visual learning for science and engineering. In *Computer Graphics and Applications, IEEE*, Sept.-Oct. 2005.
10. F. J. P. and B. K. E. Design of a haptic visualization system for people with visual impairments. In *IEEE Transactions on Rehabilitation Engineering*, volume 7, pages 372–384, 1999.
11. P. Roth, D. Richoz, and T. Pun. A multimodal system for the non-visual exploration of digital pictures. In *Interact 2003, 9th ICIP TC13 Int. Conf. on Human-Computer Interaction*, Zuerich, Switzerland, September 1-5 2003.
12. J. Sanchez and M. Saenz. 3d sound interactive environments for problem solving. In *Assets '05: Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*, pages 173–179, New York, NY, USA, 2005. ACM Press.

³ <http://www.aimatshape.net>

13. C. Sjostrom. Using haptics in computer interfaces for blind people. In *CHI '01: CHI '01 extended abstracts on Human factors in computing systems*, pages 245–246, New York, NY, USA, 2001. ACM Press.
14. S. Wall and S. Brewster. Feeling what you hear: tactile feedback for navigation of audio graphs. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1123–1132, New York, NY, USA, 2006. ACM Press.
15. S. Wall and S. Brewster. Tac-tiles: Multimodal pie charts for visually impaired users. In *NordiCHI 2006*, Oslo, Norway, 2006.
16. K. Whitelaw. Why make websites accessible?: and how? In *SIGUCCS '03: Proceedings of the 31st annual ACM SIGUCCS conference on User services*, pages 259–261, New York, NY, USA, 2003. ACM Press.
17. W. Yu and S. Brewster. Multimodal virtual reality versus printed medium in visualization for blind people. In *Assets '02: Proceedings of the fifth international ACM conference on Assistive technologies*, pages 57–64, New York, NY, USA, 2002. ACM Press.
18. M. Ziat, O. Gapenne, M. Rouze, and A. Delwarde. Recognition of different scales using a haptic sensory substitution device. In *EuroHaptics '06*, july 2006.