AUTOMATIC TRACKING ANALYSIS IN MORRIS WATER MAZE BIOMEDICAL VIDEOS

Andrés Rodríguez1, Antonio Ortega-Álvaro2, Alejandro Sola1, Juan Antonio Micó2 and Oswaldo Trelles1

1 Department of Computer Architecture, University of Málaga
Bulevard Louis Pasteur, 35, 29071 Málaga. Spain
2 Pharmacology and Neuroscience Research Group. Dept. of Neuroscience, Faculty of Medicine, University of Cádiz
Plaza Fragela, 9, 11003 Cádiz. Spain

Abstract
We present an automatic video processing system for tracking analysis in the Morris water escape testing videos. This system is able to extract automatically from such videos information (metadata) describing the spatio-temporal trajectories of animals in the maze and the timings of behavioral events such as stopping or crossing of a target area. The specific semantic metadata are produced by the system and saved in a XML file that describes the mice behaviour extracted from the video content. This description is also stored in a database using a data model that allows one to perform subsequent queries to obtain important factual and analytical information and to retrieve and visualize selected video sequences matching specific query criteria. In all cases, we have observed a significant increase in both accuracy and efficiency when undertaking such analysis using the procedures described in this paper, in comparison with current methods.

Key Words
Tracking, object recognition, video analysis, medicine.

1. Introduction
Extracting useful information from massive data files, such as those produced in the biological and biomedical area in the form of 3D images and videos, is becoming an increasingly important activity in many scientific and commercial domains. However, while the amount of image data being produced is exponentially increasing, our ability to absorb and process this information remains limited.

Manual analysis of biomedical multidimensional-moving image data allows only a small part of its visual information to be abstracted, sometimes leading to an incomplete understanding of the biological processes being observed. Vast quantities of valuable information contained in the images may be lost because of time constraints and lack of objectivity in the human interpretation of content. On the other hand, automatic procedures for recovering information from this type of video often show severe limitations and fail to detect critical events.

In a move to close this technological gap, we present in this paper an automatic tracking analysis procedure for analyzing Morris water maze video collections. The Morris water maze (MWM) [1][2] is a behavioral test developed to investigate spatial learning and memory in laboratory mice and rats. It has become one of the most frequently used laboratory tools in behavioral neuroscience. In a few words, the MWM consists in a circular pool filled with opaque water in which a small escape platform is hidden (submerged beneath the surface). When placed in the maze the animal’s task is to find the hidden platform and escape from the water. During a number of training trials, the animal learns to find the platform. By observing changes in the animal’s behavior when it is repeatedly placed in the water, the experimenter can conclude whether the animal is learning.

Although the basic procedure is pretty simple, the MWM has been used in some of the most sophisticated experiments on the neuropharmacology of spatial learning and memory to validate models for neurocognitive disorders and to evaluate possible treatments [3][4].

The animal performance during the test is recorded on a video-tape [5], since a direct analysis by an experimenter could lead to lose much behavioral information. This kind of experimental videos can be analyzed using several commercially available systems (e.g., SMART video tracking system from Panlab SL, currently in use on our lab), but usually they need special hardware systems and they have serious efficiency and accuracy limitations. Due to these limitations the current system used in our neurobehavioral lab needs continuous human calibration of parameters, supervision and by-hand corrections. These are time consuming tasks that lead to errors and subjective inconsistencies in the results.

Our proposed system solves the problem in a fully automated fashion, getting consistent and reliable results. This way, the presented system increases analytical power for uncovering and studying the effect of drugs, genetic mutations or disease on a variety of behavioral models.
2. Video Analysis

Our strategy for video analysis (see Figure 1) can be summarized in two main steps. The first one consists in the detection and tracking of the objects of interest that are present in the images and the second step is the image understanding process. Initially images must be processed and segmented to identify the discrete objects in each video frame. In these scientific videos there are two kinds of objects to detect: the maze where the animal moves and the animal itself.

Maze recognition and location is based on geometrical shape matching among the maze models and the objects extracted from the image background. This recognition will be done just once at the beginning of the analysis because neither the maze nor the camera will move on an experimental session.

Regarding the mouse, it is identified thanks to the segmentation of its distinctive color, which displays a good contrast against the light color of the maze. Tracking the movements of mice along the space/time axis of the video sequences is essential because the analysis of these movements will characterize their behavior.

As a result of the video analysis the system will generate a XML report containing the description of the observed mouse’s behavior and will also save a modified version of the analyzed video with on-screen annotated information plus an image of the complete path swum by the mouse.

The image processing strategies designed have been implemented in C++ using the OpenCV Intel Open Source Computer Vision Library [6]. In the subsequent subsections we are going into the system and methods details.

2.1 Segmentation and Tracking

2.1.1 Water maze Segmentation

The first step of our analysis approach relies on the localization of the maze. In order to locate the maze, it must first be separated from the scene background. This segmentation process is carried out using a global thresholding algorithm.

Figure 2. On the left: example video frame with a black mouse swimming around the centre of the pool. On the right: video frame histogram. Clearly bimodal, the left (darker) accumulation is the background (plus mouse) and the right (lighter) is the water maze (pool).

Thanks to the controlled experimental conditions of the video shooting, there is a good contrast between the background and the maze. The maze is a pool filled with water and with some opaque colorant added (milk for dark colored mice).

Therefore gray level histograms of the video frames are clearly bimodal, and the two main accumulations (background and maze respectively) are well separate (see histogram in Figure 2). The mode method [7] is used for the design of the algorithm for automatic threshold detection. The approach implemented finds the two highest local maxima on the histogram first, and then the threshold is detected as a minimum between them.

Once the scene has been segmented (see Figure 3), the algorithm iterates over every discovered connected region until the maze is matched, recognized and characterized. The maze is recognized based on its 2D shape. For every region in the scene the algorithm checks whether it has the maze shape.
The signature of contours [8] is used for characterizing and detecting the maze shape. The signature is a functional representation of a contour, which may be generated in several ways, but in this case distance-versus-angle signature [9] is used. This signature function plots the distance from the geometric center of the contour to its boundary as a function of angle.

Figure 3. On the left: segmentation of one video frame based on mode method thresholding. On the right: contour of the connected region that belongs to the Morris water maze.

The distance-versus-angle signature is the result of a transformation from Cartesian to the polar coordinate system. Using the geometric centre of the shape as the origin of coordinates guarantees that such a transformation is invariant to translation. Additionally, the distances are normalized with respect to maximum, so that transformation is also scale invariant. However, it is still dependent on the orientation of the original shape. There are several suggested ways for selecting a starting point (angle) for generating a signature invariant to rotation [8], but instead of trying these, the invariability is handled by rotating one of the shapes over the other until finding the best correspondence. The rotation in Cartesian system become a translation on the angle axe in the polar coordinate system, therefore the comparison procedure is simple and fast.

An example of the signature of the maze contour shown in Figure 3 is plotted in Figure 4. The system stores the signature expected for a model maze, so both signature functions (observed and model) are compared using correlation to decide whether the observed object is the maze or not. In order to input a valid model maze signature, our system allows using a video or image example for acquiring the correct maze signature calculated from the contour pointed out in the segmented image manually by the user.

As can be seen on Figure 3 there are some artifacts in the detection of the contour of the Maze, which should be round in shape. These artifacts are originated by four cues located on the edge of the maze (these cues are designs drawn on paper and hung on the edge of the maze, visible to the animal). These designs serve to the animals as spatial references to remember the location of the escape platform. It is an easy task to remove the errors in the contour by the extrapolation of valid radius on the contour based on the surrounding correct contour points (see Figure 5).

Figure 5. Maze contour correction. Left: original contour obtained by image segmentation. Right: final contour after elimination of cue panels artifacts.

Once the maze have been recognized, by matching the contour that satisfies the correlation test with the maze model signature, the maze is characterized by means of its position (geometrical center), size (mean diameter) and the signature of the contour (necessary in case of contour is not perfectly circular).

Regarding the orientation of the maze, we suppose the north of the maze match the north orientation of the camera viewing field. Finally we automatically divide the maze in four quadrants (NW, NE, SW, SE), and we also define some ROIs of interest for further analysis: the periphery of the maze arena, the center of the maze (total minus periphery), the escape platform and a circular critical zone around this platform. There are also two important reference points (RP): the center of the maze, and the center of the platform. Most variables included in the tracking analysis were related to these ROIs and RPs.

The position of both the maze and the camera are constant along one experimental session, therefore its recognition and localization can be done only once at the beginning of the video sequence.

2.1.2 Mouse Tracking

The next step in the video analysis is to locate the mouse inside the maze. In the videos analyzed the mouse is black, and then its image in the video frames will be
composed by the darkest pixels inside the maze. The particular mouse characteristic color is exploited for its identification using a p-tile thresholding \[7\]. The relative size of the mouse with respect to the maze size is used as a first approach to estimate the mouse size in pixels. Then, based on the image histogram, a threshold can be chosen such that only mouse pixels have gray values lower/darker than threshold. In Figure 6 the results of this image segmentation algorithm that leads in the localization of the mouse are shown.

As we have mention above, experimental mice are “usually” black, but this is not always true. There are also experimental mouse strains that are white. To be able to segment white mice inside the water maze, some kind of water-soluble dark dye should be added to the water in the pool. Then the mouse would be composed by the brightest pixels in the image and it can be segmented using the same p-tile thresholding approach discussed above for black mice. In this case a threshold is chosen such that only mouse pixels have gray values higher than threshold.

\[Figure 6.\] On the upper left: original video frame with the black mouse swimming in the centre of the maze. On the upper right: the image after the background subtraction. On the bottom: segmentation of the mouse (pointed by an arrow) using p-tile thresholding.

We measure some basic parameters describing the mouse by analyzing the connected region of pixels that form its image in the video frame. This information is constituted by: size (in pixels), location (mass center), orientation angle (principal axe) and shape (length and width respect principal axe). The mouse moves around the maze along the time axis of the video so the localization of the mouse must be done every frame of the video for tracking the path followed by the animal (see track example in Figure 7).

Some heuristic tests are conducted for checking the correct tracking of the animal: the position of the mouse must be continuous in space and time, the size of the mouse is limited inside an interval, the mouse can not disappear or multiply, etc. If some of these tracking tests fail due to noise or artifacts in the source image, then the position of mouse is not updated until a correct status is promptly reached again. This way the system solves the issue of tracking even if there is moderate noise in the image (e.g., due to low quality tape recordings).

\[Figure 7.\] Complete example track of mouse starting at north and heading the escape platform (after three loops) in the SE quadrant. The track is drawn in real-time during analysis and it is also saved to disk for archive and further access.

\[2.2\] Semantic Analysis of Mouse Actions

More elaborated information describing the action in the video must be extracted from the localization of maze and tracking of the mouse movements. We must decide and describe what, when and where is the mouse doing and record the significant occurrences or events taking place in the video action for a subsequent analysis. The automatic detection of specific behaviors and interactions in the video will allow the synthesis of video content descriptions of high information value.

From the neurobehavioral point of view, the most important events that the system must recognize are the mouse entering and leaving the ROIs defined in the maze. Taking into account the position of the mouse and the geometrical location of every ROI of the maze, the system guesses the relative position of the mouse inside the maze and assigns one or more states to it:

- In the center of the maze
- In the periphery
- In the critical zone around the platform
- In the escape platform
- In one of the four quadrants (NW, NE, SW, SE).

Based on these states the system records the log of visited ROIs, annotating the time spent in the ROI and the distance covered in the exploration, the mean velocity, the turning rate (angular velocity).

While the analysis is running the system displays the video with on-screen real-time information. One of the screen-shots of the video analysis application is shown in Figure 8, where the mouse position and orientation is
denoted in real-time by a circumference crossed by an axe. The maze ROIs and quadrants structure is also outlined. This annotated video is also stored for later retrieve and exploration. The track of the animal is accumulated and displayed on other window also in real time (see Figure 7), offering more information to the scientists during analysis.

Figure 8. On-screen annotated output frame displayed and stored during video analysis. The position and orientation of both maze and mouse are denoted. Information regarding the time spent and the position of the mouse (ROIs and quadrants) is shown in real time.

At the end of the video analysis a XML content report is generated in a file describing the mouse behavior and the observed events.

3. Video Content Description Database

The experiments we are analyzing in this work involve the observation of a number of different animal groups through successive trials. After these observations all the semantic content information extracted from the life experiments or the video recordings require further in-depth exploratory and statistical analysis. The neuroscientist looks for patterns in the animal learning process and correlations between this learning and various drug administration or other experimental conditions.

With this work we want to address the whole process of video content analysis and not only its automatic extraction. Therefore a database model has been designed to store the semantic content description of the videos analyzed. The video sequences themselves are also stored in the same database together with their content descriptions.

A prototype application for managing this database has been developed, allowing: (a) to register and store new analyzed videos and its content description in the database, (b) to search in the database allowing video-content-based queries that retrieve and visualize the videos or video segments of interests and the content descriptors associated with them, (c) to analyze a collection of video track descriptions to find patterns, trends and relationships between different dependent variables (e.g. find associations between learning patterns and diseases or drug treatments).

The content descriptors (also called video metadata descriptors [10]) are organized in the database following an organization model for scientific video content description [11][12] that can be applied in this case as shown in previous work [13]. Experiments analyzed by the system are automatically inserted in the database, registering the video file and the scenes that compose the video (the media entities). The mice that appear in the scenes are registered together with the behavioral report obtained from the analysis process. All events of interest (crossing of a ROI) are also registered in the database, indicating which mouse participates in and the time coordinates inside a scene where it takes place.

With all this video content information available and accessible in the database, the researcher can locate and retrieve any video sequence (or segment of it) based on its content. This retrieving and visualization tool enables the true exploitation of the video analysis system analytical power.

An example of graphical visualization of results of our prototype database application is in Figure 9, where the tracks of a sequence of four trials for the same animal are plotted. The experimentalist can have at a glance a good impression of the learning progress of a specimen. Other good example is shown in Figure 10, where learning abilities of two different mouse strands are compared using the time spent to escape from water as main discriminant variable.

Figure 9. Sequence of successive tracks of the same animal in the course of four learning sessions. Shown tracks start from south and the target escape platform is in the south-east quadrant.
**4. Conclusion**

The video analysis system presented in this work enables the automatic analysis and semantic description of the content of neurobehavioral research videos featuring Morris water escape task (MWM). The system produces meaningful specific metadata describing the semantic content of the videos that support subsequent queries to obtain important factual and analytical information, and to retrieve selected video sequences matching the query criteria.

The obtained results have been validated by alternative conventional methods of analysis, and significant increases in both accuracy and efficiency have been observed. The system gets consistent and reliable results based on its automated operation in the light of detailed application domain knowledge. The system will be a powerful tool for the analysis of mouse learning research videos and its application represents a contribution towards the study of new drugs for the treatment of cognitive deficit in neurological and mental disorders.

At present our prototype system is being used to extract knowledge in an automated manner from a range of medical research videos, it is clear that this system architecture could have widespread usefulness for the analysis of other types of moving image data in the scientific application domain.

**5. Acknowledgement**

This work has been partially supported by EU grant QLK3-2001-01473 (sub-programme area “Quality of Life and Management of Living Resources” - Key Action “The Cell factory”), Ministerio de Sanidad grant FIS 01/1055 and FIS PI031430 and Plan Andaluz de Investigación grant CTS-510.

**References**


