VidART - A TOOLBOX FOR AUTOMATIC HUMAN POSE AND GESTURE RECOGNITION

Automatic annotation of human behavior

State-of-the-art computer vision techniques can replace the need of tedious manual annotation in human behavior research by an automatic analysis tool.

Analyzing human behavior is complex and manual annotation of the desired behaviors is tedious. To help human behavior researchers, in this Knowledge Transfer Project (KTP) the Video Analysis and Recognition Toolbox (VidART) is developed, which is a toolbox consisting of computer vision techniques to automatically detect and track people in the scene and identify human poses and gestures from videos.

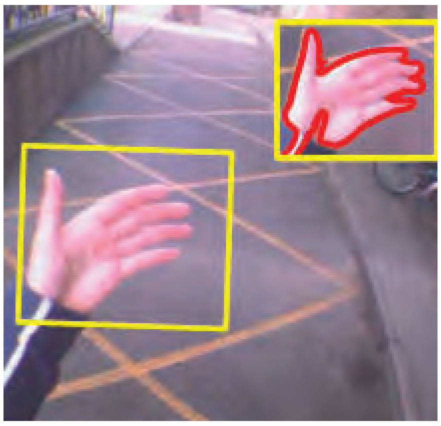
Noldus Information Technology has developed software to manually annotate and analyze human behavior from video footage. Automation recognition of such behavior can speed up the annotation and analysis process significantly. Before any type of behavior can be recognized, the system must detect the subjects in the scene and track them. An intuitive way of detection in an indoor environment is background subtraction, where the current frame is compared to an image or model of the background. Because we assume multiple cameras are available, we investigated a multi-view background subtraction technique, where pixel-to-pixel correspondences between two images can only be distorted by a foreground object in the scene. After the detection, the subject must be tracked over time. The idea behind our real-time contour tracker, is that we first use an efficient rigid body tracker to restrict the region of interest, and next, have the computationally intensive part of the tracker find deformations in the contour. Although this is a single-view object tracker, these contours can be combined to a 3D voxel reconstruction when multiple cameras are available. When subjects are separated in space, pose estimation can be performed on each person individually. This results in positions of each body part, like the head, hands, etc. To recognize pre-defined gestures of a specific subject, the movements of a (combination of) body part(s) should be identified and interpreted. Therefore, estimating the body pose is a crucial step.

Restaurant of the Future

Before any software, like VidART, can replace an existing way of analysis, a representative test must show that the software performs equally well in accuracy and robustness. Since Noldus IT is one of the research partners participating in the Restaurant of the Future, we use this facility in this KTP as our application domain. This restaurant is equipped with pan-zoom-tilt cameras, which allows for close observation of consumer eating and drinking behavior. The restaurant is an indoor environment, but not controlled (large windows, non-static background, etc.), the subjects act naturally and their appearance is completely unconstrained. This makes this facility a suitable test environment to show how the software works in a real-life situation.

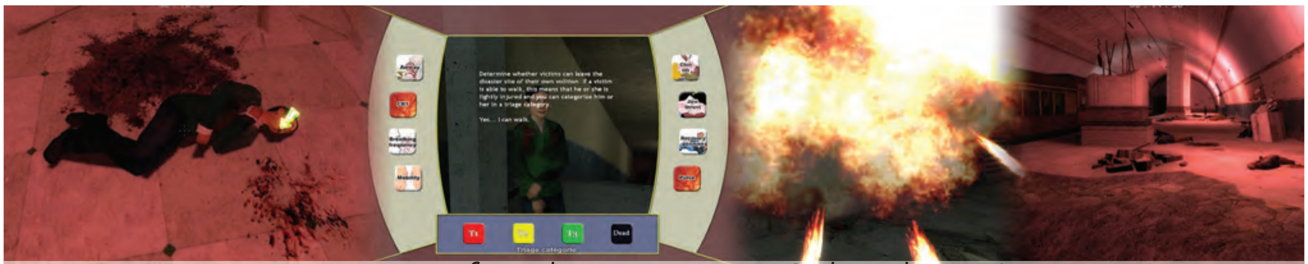
Utrecht Multi-Person Motion (UMPM) benchmark

Since the Restaurant of the Future is a complex environment, we created the UMPM benchmark for multi-person pose estimation and gesture recognition validation using a marker-based motion capture system available at Utrecht University. Each scenario of the benchmark concerns multi-person motion and interaction in a controlled environment, includes 1 to 4 subjects visible in the scene, and maximally two of them wear reflective markers to measure joint positions of the body. The joint positions found by the pose estimation software can be evaluated with these measured ones. This benchmark is available via http://www.projects.science.uu.nl/umpm/ to the research community to advance the field of multi-person human motion capture.



Nico van der Aa is computer vision researcher at Noldus Information Technology BV. After obtaining his Ph.D. title in numerical mathematics, he became involved in image processing and computer vision techniques. During the GATE project he actively participated in the Multimedia & Geometry group at the department of Information and Computing Sciences at Utrecht University.
The VidART project is supervised by Remco Veltkamp from Utrecht University and Herman Roosken from Noldus Information Technology BV.

Game research
for training and
entertainment



Impressions from the serious game 'Code Red Triage'.

In the GATE research theme Learning with Simulated Worlds we study how to best create games for learning. We study adaptivity of games, various learning paradigms, design rules for educational games, and how the abilities learned in games transfer to the real world. Indeed, advances in building worlds, populating worlds with virtual characters, and interacting with virtual worlds (other research themes in GATE), cause real and simulated worlds to merge, which opens new ways for education, for both teaching and learning. Gaming is often experienced as fun, and thus may be a vehicle to self-driven and highly motivated learning. These controlled virtual environments can provide decision support and situation awareness support. Combining virtual gaming and real environments into augmented reality takes the potential for learning even further.

For example, we investigate how elements of the game such as virtual characters can adapt themselves in a natural way to the level of the user. We have designed a software framework for the Game Adaptation Model that can be used to orchestrate the adaptation of individual agents within a game. There are a few elements that play a key role in this model. First, all the agents behave according to their role specified in the game model. The role specification puts some limitations on the extent to which agents can adapt their behavior, because the behavior has to comply to the general storyline, interaction norms, etc. of the game. Secondly, the adaptation engine will centrally register the need for adapting the difficulty level of the game and based on the result put out a call for adaptation to the agents. Based on the bids of the agents a suitable combination of adaptation is determined, which is subsequently performed by the agents. An important aspect in this process is that the agents determine the ways they might adapt based on their current status and goal(s). This ensures that agents always adapt in a natural way, fitting with their current and past behavior. In this way the model ensures a balance between global and local control of the adaptation where both local and global consistency is preserved.

We also perform research to establish design rules for learning, specifically on the gaming 'dispositif', the properties of persuasive game design, and the role of narratives in games. We developed a theoretical model that offers an alternative to existing academic approaches to game narratives. The model explicates the logics behind two primary ways in which avatar-based 3D games deal with stories. One of these logics focuses on players as implied authors who guide heroes through challenging trials and tribulations, and intervene in their faith by controlling them. The other logic focuses on players as embodied participants in the story world; players become the hero and have adventures of their own. Drawing on theories from game, film, theatre and communication studies, the model defines these two logics, presents their characteristics and explains potential problems when the logics co-exist in the design of one and the same game.

We also experimented on how cognitive design principles can be used to improve learning with a serious game. To this end we created a serious game, Code Red Triage (see picture above), and systematically varied elements in the game design to measure the effects on learning and engagement. We investigated the ability of auditory and visual cues to guide the player's attention, the way in which instruction has to be embedded into the game, how curiosity can be engendered; and discovered, among other things, that games can be made significantly more efficient when they adapt to the player's performance and that incorporating surprising events at key moments leads to the player constructing superior knowledge structures. The interaction between game design and the player's cognition can furthermore be described with "GameDNA" (Game Discourse Notation and Analysis), a notation tool we are developing that should benefit game designers in creating better serious games.

Equally important is the transfer of gaming, the rate and efficiency with which learned skills are transferred to practical situations. We conducted an experiment in the F16 flight simulator to determine the amount and kind of transfer of training of two flight games (Falcon 4.0 and Microsoft Flight Simulator) to professional F-16 flight skills. This experiment was aimed at determining the amount and kind of transfer of two flight games and non-flight games to professional F-16 flight skills. The tasks these subjects had to perform in the F-16 simulator were chosen in collaboration with a very experienced ex F-16 test pilot of the Royal Netherlands Air force. The results show that Falcon 4.0 gamers perform significantly better on all the tasks compared to the Microsoft Flight Simulator gamers and the non-flight gamers. It can therefore be concluded that Falcon 4.0 gamers acquire more skills and competences compared to Microsoft Flight simulator gamers for professional F-16 flight skills.

Game adaptation, design rules, cognition-based principles, and transfer of gaming are all aspects that are of paramount importance in game-based education. In all these cases, further innovation is needed. More knowledge and insights are necessary to further increase the effectiveness of game-based learning and teaching.

Remco Veltkamp Utrecht University