



**Universidade de São Paulo**

**Biblioteca Digital da Produção Intelectual - BDPI**

---

Departamento de Ciências de Computação - ICMC/SCC

Comunicações em Eventos - ICMC/SCC

---

2015-10

# Mobile video annotations: a case study on supporting rehabilitation exercises

---

Brazilian Symposium on Multimedia and the Web, 21th, 2015, Manaus.

<http://www.producao.usp.br/handle/BDPI/49619>

*Downloaded from: Biblioteca Digital da Produção Intelectual - BDPI, Universidade de São Paulo*

# Mobile Video Annotations: A Case Study on Supporting Rehabilitation Exercises

Bruna C. R. Cunha, Rodolfo Dias Correia, Maria da Graça Pimentel  
Universidade de São Paulo  
Instituto de Ciências Matemáticas e de Computação  
São Carlos, SP - Brazil  
{brunaru, rodolfod, mgp}@icmc.usp.br

## ABSTRACT

Annotations enable us to highlight key points or add information to content presented, for instance, on paper or digital media. Even though smartphones and tablets facilitate video capturing, currently only few applications allow limited video annotations using the mobile device itself. Given that video annotations can assist many tasks that depend or can be enhanced by video capturing, in previous work we have contributed with a tool for allowing multimodal video annotation using mobile devices. Upon experimenting with the tool, we identified that reviewing rehabilitation exercise videos can be especially enhanced with video annotations. Professionals in the field of physiotherapy and occupational therapy can add relevant annotations for their patients to improve performance of their exercises. After an evaluation with a specialist in occupational therapy, we identified new requirements associated with the monitoring of patients. We then identified the opportunity to develop a monitoring system with the collaboration of two rehabilitation specialists. Since these two specialists work with mirror therapy, we defined requirements that are relevant for this type of therapy. The system was developed to support the monitoring of exercises combined with video recording and annotation capabilities. The system aims to support rehabilitation therapy by distance: therapists can monitor patients as they record videos that are sent for evaluation. We identified requirements that may be applied in many scenarios, however mirror therapy demanded many specific requirements related to the recording of videos.

## Keywords

mobile capture; rehabilitation; usability; mirror therapy; design.

## Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentations]: User Interfaces—*Graphical user interfaces (GUI)*; H.5.1 [Information Interfaces and Presentations]: Multimedia Information Systems—*Video*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

*WebMedia'15*, October 27–30, 2015, Manaus, Brazil.

© 2015 ACM. ISBN 978-1-4503-3959-9/15/10 ...\$15.00.

DOI: <http://dx.doi.org/10.1145/2820426.2820449>.

## 1. INTRODUCTION

Nowadays smartphones and tablets are highly popular and accessible to the population in general. These devices enable a rich user interaction and come generally equipped with good quality cameras. These facts have stimulated the production and distribution of multimedia content by users. In addition to producing content directly, users can add annotations on different types of media using their mobile devices. Annotation capabilities on textual documents are practically mandatory in any competitive document viewer. On images, annotations are also increasingly popular. However, although there are tools that allow annotation of texts and images to be directly made via mobile devices, video annotation applications are rare. Given that video annotations can assist many tasks that depend or can be enhanced by video capturing, in previous work we have contributed with a tool for allowing multimodal video annotation using mobile devices [13].

Observing that video annotations are useful in different scenarios, in this paper we discuss the opportunity of providing video recording and annotation of rehabilitation exercises. Exercises are of great importance in rehabilitation of individuals with physical or cognitive disabilities [29]. Different areas of computer science have contributed with diverse systems and applications to assist in performing rehabilitation exercises, e.g. robotics [24], virtual reality [20] and sensors [9]. However, to the best of our knowledge video annotations on mobile devices as a support tool for performing exercises have not been explored in rehabilitation area.

Professionals from the fields of physiotherapy, occupational therapy, phonoaudiology, among others, plan personalized physical, motor or vocal exercise routines for each patient. Even though patients usually need to repeat the planned exercises several times everyday, in most cases professionals cannot supervise them on a daily basis. This implies that patients have to perform self-treatment at home between consultations [26]. Furthermore, for a treatment to be effective patients must repeat the exercises correctly [28]. Since learning these exercises depends a lot on observation and memory, patients are prone to forget how to repeat them accurately. In the context of unsupervised exercises, researchers show that both young and old patients forget and incorrectly perform physiotherapy exercises – they also show that older people forget more than young people [28]. Considering these issues, some professionals adopted the use of video recordings to support their patients' exercise routines. Miller et al. [26], for example, demonstrate that the use of videotaped exercises is useful for physiotherapy patients and that it serves as a support tool for their treatment by helping patients to complete the prescribed exercises accurately. Similarly, Lee and Connor [31] reported that videotaped exercises increases adherence of patients who need to perform voice therapy. These

scenarios motivate the use of video-based technologies toward supporting rehabilitation exercises.

To understand the use of mobile video annotations in the rehabilitation exercises scenario, we firstly interviewed one specialist of occupational therapy that already records videos of her patients during sessions. Her aim is that patients can review their own videos when necessary and thus reproduce the exercises more accurately. During sessions, the specialist uses her patients' smartphones to record in video the prescribed exercises. The specialist stated that annotations would be of great help: she could enhance the videos by adding important information so her patients could improve the way they perform their exercises. Furthermore, patients could recording their exercises by themselves, so therapists could watch videos to evaluate and add annotations when necessary. As a result, we extended our video annotation tool to include a video capture option so it could be used for recording and annotating exercise sessions. We then invited the specialist for an evaluation of the tool and for contributing with new ideas: results indicate a good potential of the application in the proposed scenario and provided new requirements associated with the monitoring of patients. Giving the demand for monitoring functions, in collaboration with two rehabilitation specialists we developed two applications: one for therapists and another for patients. These specialists are working extensively with mirror therapy [16], thus we focused in this therapy. This experience helped us to understand some of the requirements of monitoring systems for rehabilitation. We describe the main requirements we found in this paper.

The remainder of the paper is organized as follows: Section 2 analyzes the related work on mobile video annotation and video annotation applications. Section 3 details our mobile video annotation tool. We describe our evaluations and discuss the results in Section 4. Section 5 describes our monitoring applications and their requirements. Section 6 presents our concluding remarks.

## 2. RELATED WORK

Ubiquitous computing is about putting computing in the real world. Over the past few years, current mobile technologies have made creation of solutions for multi-disciplinary problem domains more affordable [1]. Many works are proof of how mobile devices can support people and solve different problems of their lives. For example, [7] developed a set of PDA applications to support psychotherapy processes and to help therapists and patients to carry their therapies more efficiently. [3] presented a mobile application that helps blind people to travel more independently and safely. As another example, [15] created a set of tools for the self-management of people with Parkinson's disease that also allows doctors to follow their patients' symptoms and medication adherence.

Given that mobile devices such as smartphones and tablets were released recently (the first iPhone, for example, was released in 2007), earlier research in the area of mobile-based video annotation explored rather limited devices in terms of mobility and interactivity. The StuPad, for example, allowed pen-based annotation to be carried out on slides from live lectures and automatically synchronized with a corresponding video [30]. M4Note tool uses tablet PCs to capture voice and pen-based interactions as annotations. The tool allows the annotations to be added during the capture or the playback of a video [19]. Also for tablet PCs, the Watch-and-Comment Tool allows to add text, audio and digital ink annotations on a recorded video. Furthermore, collaborative features are present such as an instant messenger and sharing of annotation files by a Peer-to-Peer network [8].

Research reporting the study annotations on feature phones are restricted and specific due to the lack of resources of such devices.

For instance, The Mobicon tool enables the annotation of videos with automatically generated metadata, user can accept a recommended annotation but cannot add free comments [22]. The Mesh tool, on the other hand, enables users to assign keywords and textual annotations to specific moments of a video using the numeric keypad [10].

In contrast, mobile devices like smartphones and tablets allow a richer interaction, enabling developers to design complex applications. However, despite the maturity of current mobile devices, there are not many research works involving video annotations. [21] reported an Android application for tablets that allows text and digital ink annotations on a video transmission. The tool also has collaborative capabilities that allow users to see annotations made by others. However, the focus of the work is a framework for collaborative annotation. There is no investigation about the tool's user interface and details about its functionalities. [25] also presented an Android application for tablets that allows the addition annotations over a recorded video. It is possible to add text, digital ink and audio annotations, but audio is stored as text using a speech-to-text functionality. The tool shows marks on the timeline for each annotation to support navigation. Although there are concerns about the user interface, the authors did not conduct evaluation experiments. These two Android applications were designed for tablets. There is no discussion about tests on smartphones and if the design would fit well on small screens. None presents integrated capture capabilities. Furthermore, there is no evaluation with users and no discussion around its peculiarities in terms of user interaction.

Research have been reported towards applying video annotation in several domains, as in the case of annotating endoscopy videos [25] of family videos [2]. As another example, [4] present a PDA tool to annotate emergency videos using pre-defined graphics and labels. The Creation-Tool is a multimodal video annotator designed to assist the creative processes of choreographers [6]. The EVA platform is a web video annotation tool that aims to help asynchronous video teaching and learning by using temporal text annotations and navigation functionalities [23]. The VidWiki tool was conceived to enable users to collaboratively annotate educational videos to correct errors and improve legibility [11]. Although there are many similar studies in different fields, we could not find works on video annotation as a support tool for rehabilitation exercises.

## 3. ANNOTATION TOOL

In previous work we contributed with a video annotation application for Android tablets and smartphones [13]. The application, named MoViA Tool<sup>1</sup>, is executable from version 2.3 of Android to the newest 5.0 version. The application has a video capture functionality that allows the adding of annotations as temporal marks to the video. The application also allows multimodal annotations to be added to pre-recorded videos. The annotations are time-dependent, meaning they are connected to specific moments of execution of a video. During video playback, the system allows users to add text, electronic ink and audio annotations. In the next sections, we discuss details of the application development, interface and features.

### 3.1 Development

The developed tool went through four development stages and three evaluations. The requirements for the first version of the tool were defined after a literature review considering state-of-the art related work. This approach was chosen so we could offer users a working prototype when collaborating with us in the identification of

<sup>1</sup><https://play.google.com/store/apps/details?id=br.usp.icmc.movia>

requirements. The first version only had textual annotation and navigation capabilities. We presented the first version to potential users who participated in a usability test. The results from this cooperation with users were used to correct issues and implement new features, which led to a second version of the tool. The main innovations of the second version were multimodal annotation and sharing functionalities. We then conducted a heuristic evaluation that counted on the collaboration of ten computer scientists. The results from the heuristic evaluation were combined with those of a systematic review in collaborative annotation systems to define new requirements for the third version. The third version went through an collaborative evaluation with an occupational therapy specialist, as described in Section 4.2. We solved the issues identified in this last evaluation for the current version of the tool, presented in this section. Evaluations are described in Section 4.

### 3.2 Annotations

In the application, the first screen presents options for capturing a new video, annotating an existing video and importing annotations. If the user chooses to create a new video, the video capture screen appears. The application has its own implementation of video recording that allows adding temporal marks during the capture. While capturing, users can press the mark button in moments of interest or in situations they want to add more information. After the video recording, the tool generates a document with textual annotations containing the selected points. When reproducing the video in the application, users can navigate by the marked points and add new annotations.

Figure 1 presents the video annotation screen. The vertical layout is called edit mode. In the edit mode, users can add annotations, navigate, visualize contextual information and use sharing options. We observed that it is easier to type on smartphones in portrait mode. In landscape mode (called display mode) the application plays a video and its annotations on full screen without allowing the addition of new annotations.

The tool calculates the display duration of textual annotations using a formula that considers the average of words read per minute on mobile devices. The formula for estimating the duration is given by  $d = n/(wm/60)+3$ , where  $d$  is the duration in seconds,  $n$  is the number of words in the text and  $wm$  the number of words read per minute. The value of words per minute is fixed and equal to 200. We chose this value based on averages reported by studies that investigate the reading speed on mobile devices [18]. The additional time of 3 seconds is for displaying the text for longer than the reading time, considering that annotations need more time than simple subtitles.

For adding electronic ink and audio annotations, users must click on the buttons that enable them. By enabling the audio annotations, the video is paused and the audio recording starts. It is required to pause the video while audio recording because the audio from the video may cause interference. The video is paused again during playback of an audio annotation. The video is also paused for adding an ink annotation, but ink annotations can be made and is shown without the need of pausing the video.

### 3.3 Navigation

The application generates the navigation screen automatically based on existing annotations. For each available annotation, the screen shows a video thumbnail of the annotated instant along with the time of occurrence, the annotation text and a button for editing or removing the annotation. If the thumbnail or the text area is touched, the progress of the video is set to the time of the annotation, i.e., it jumps to the particular point of interest. This strategy

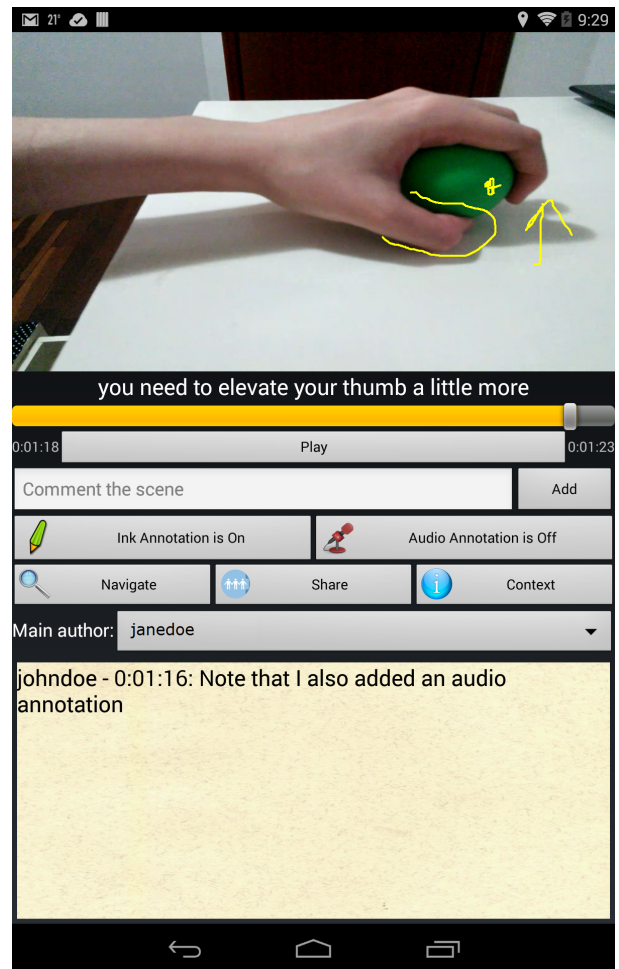


Figure 1: Playback and annotation screen in edit mode.

intends to improve the search and selection of points of interest to select them accurately in touchscreens, which is usually complicated due to issues related to the precision of the fingertips. The navigation screen for audio and electronic ink follows the same model and are selectable via the radio buttons at the top of the navigation screen.

### 3.4 Sharing

Our application allows users to share their annotations and videos using the device's sharing options. During the reproduction, textual annotations of multiple authors can be displayed simultaneously in a reserved area (bottom area on Figure 1). The tool does not simultaneously play annotations in ink and audio originated from different users. The reason is to avoid overlapping conflicts of concurrent annotations. However, the application enables dynamically changing the playable annotations during reproduction. To switch between annotations of different authors it is necessary to set the main author. Users can select the main author using a combo box that lists all possible authors based on available annotation documents for the current video. Thus, the tool reproduces digital ink and audio annotations according to the selected author. At the same time, textual annotations from the main author are displayed just below the video.

### 3.5 Context Information

The captured context information aims to assist users and researchers. For example, for users, location, event and time can help to identify or remember the circumstances of a captured video. For researchers, detailed information about used devices is important to help them to identify issues related to specific devices. All the information is captured without the need of explicit user input. The dimensions and the captured information are the following: who (user responsible for the annotation), what (media type and event from agenda), when (creation date of annotation documents, where (physical location), how (device model and OS version).

## 4. EVALUATIONS

We conducted three evaluations on the tool: a usability evaluation with users, a heuristic evaluation with specialists and a user evaluation with an occupational therapy specialist. In this paper, we focus on describing the evaluation conducted in collaboration with an occupational therapy specialist. We first briefly describe the evaluations conducted earlier (detailed results are reported elsewhere [12, 14]).

### 4.1 Previous Evaluations

For the usability evaluation conducted on the first version of the tool, we counted with the collaboration of six users. For the heuristic evaluation, we had ten computer scientists, that had already completed a HCI course, as evaluators. These two evaluations heavily contributed with the identification of usability problems and requirements. We list the most relevant of them with our proposed solutions:

- Display time of textual annotations: estimate display time according to the number of words read per minute.
- Touch accuracy on labels (annotation text on navigation screen): not only the label but area around it should be touchable to increase accuracy.
- Video selection: video files should be presented with a picture to assist recall.
- User creation: device's Google user account should be used when possible.
- Feedback messages (Android Toast) not noticed by users: when this happens, messages should be shown at center of screen.
- Activation buttons (audio and ink buttons): user should be alerted about function activations by multiple feedbacks.
- Automatic pauses: media should be automatically paused when adding annotations.
- Audio conflict: pause video or lower its volume while audio recording.

### 4.2 Evaluation with an Occupational Therapist

The specialist in occupational therapy, who participated in our interview and evaluation, works with patients who need to perform motor rehabilitation exercises. In this sense, the work of the professional is to propose personalized exercise routines. The patients should repeat these exercises daily, but in most cases, the presence of a professional during all exercise sessions is not possible. Therefore, patients should remember the exercises and repeat them without the guidance of a professional. This task is highly influenced by memory and it is common for patients to forget or incorrectly perform the exercises. To support this task and help her patients to remember how to perform the exercises, the specialist uses her

patients' smartphones to record videos while they carry out exercises. Thus, patients have their prescribed exercises recorded on their own devices being able to review them if they forget or are not sure about how to perform them.

The evaluation followed the Cooperative Usability Testing (CUT) protocol proposed by [17]. A CUT is divided in interaction and interpretation sessions. During the interaction session, a test user tries out to application to identify usability problems while evaluators mainly observe. In the interpretation session, the test user and the evaluators join their expertise to discuss and understand encountered problems and future requirements. In this experiment, we used a Nexus 7 (2012). We firstly introduced the tool explaining its features. Then, we let the specialist test the tool, using all features, while observed by us. In the end, we discussed her experience with the tool, the usability problems and what kind of functionalities she thinks that would support her work even more. One limitation of this evaluation was that it was not conducted at the professional workplace because this choice would bring additional obstacles, such as the fact that it would involve third parties consent (patients). Next, we present the results of our interpretation session in terms of the specialist opinion about the features, encountered usability problems and new requirements for the tool.

According to the occupational therapy specialist, the use of annotations on exercise videos recorded during patients' sessions is of great use. After recording a video, she could enrich it with relevant information for her patient. For instance, she could use text or audio to record recommendations and ink to highlight specific points on the video. Another important observation that surfaced is that patients might also record exercise videos by themselves and send them to their therapists to analyze and add annotations. The sharing functionality is essential. It is important to be able to easily share videos and annotations so that therapists and patients have a copy of the enriched session. The navigation functionality helps the professional to find and modify marked points. Navigation should be very useful for patients: they do not need to watch the entire video or randomly seek in it to review their therapist's comments or find an specific exercise. Finally, the context information can help users to recall details of the session, for instance, when and where the session was held.

With the specialist collaboration, we could also identify a problem. The specialist recorded a video in portrait mode but the tool incorrectly calculated the proportions for vertical videos, causing a distortion during its playback. We fixed this issue in the current version of the tool.

In terms of new requirements that would be useful for the therapist, we highlight two main ones: marking video while capturing and patients monitoring features. During video recording professionals can observe specific moments that they want to annotate later. To facilitate this task, we designed a solution that allows users to mark points of interest when capturing a video. After capturing, users can easily access and edit these points on the navigation screen. The specialist also confirmed, upon our prompt, that she would like the application to provide monitoring features. In addition to recording and annotating sessions, she would be interested in being aware of her patients' exercises routine. For instance, she wants to know if her patients successfully performed their prescribed exercises at scheduled times or if they had issues that made them to quit. The application could also send alerts to patients, helping them remember schedules of their exercises. If integrated with video recording and annotation capabilities, a monitoring system could also work as an distance evaluation environment. Along with monitoring features, patients could record videos of their exercises and send to professionals for annotation. In this way ther-

apists could monitor not only if patients are following their plan, but also observe their evolution and support them by distance. We believe that monitoring features can improve treatment since therapists would be able to identify issues and also increase exercises adherence. We observed that other works about rehabilitation exercises also recognize demands of solutions for patients monitoring. In their work, [5] identified that professionals are not able to correctly estimate adherence to unsupervised exercises of their patients. This implies a great difficulty in identifying whether prescribed exercises are ineffective or patients are non-adherents and need additional support to perform their exercises. The authors also state that there are currently no cheap and easily available solutions to this problem and they believe in the potential of mobile applications. Considering these results, we developed a monitoring system that integrated our video annotation tool. This work is described in Section 5.

## 5. APPLICATIONS FOR MONITORING REHABILITATION EXERCISES

To investigate monitoring functionalities for rehabilitation exercises we developed a system with the collaboration of two occupational therapists. These two therapists are also researchers that experiment new techniques and therapies with patients. More specifically, they are working with mirror therapy and prescribed home-based exercises for rehabilitation of hand movements. The mirror therapy consists in perform exercises with one limb while the other limb is hidden by a mirror. This treatment was initially used to treat phantom pains but is used nowadays in neurorehabilitation [27]. Mirror therapy consists in perform exercises with a healthy limb while the affected limb is hidden by a mirror. Patients must see the healthy limb reflected on the mirror during exercising since the goal is to induce illusory perceptions [16]. These therapy scenarios generated many specific requirements for our monitoring system. Our system is composed by two mobile applications, one for patients and one for therapists, we describe them separately. All requirements were implemented as requested by therapists and the applications are fully functional.

### 5.1 Application for Therapists

The application developed for the therapists is essentially about manage patients, prescribe exercises and evaluate videos recorded by patients. The application was named Interactive Occupational Therapy for Therapists and its main requirements are described as follow.

- Manage patients: therapists must be able to add new patients and consult details about patients' personal data and their clinical conditions such as affected hand, dates of lesions or surgeries.
- Personal messages: therapists expect to trade private messages with their patients inside the application. Therapists and patients must be able to communicate without relying on external message tools.
- Manage sessions: therapists plan personalized exercise sessions for each patient. Sessions are continuously modified since therapists need to adapt them according patients' improvement during treatment. A session must contain which exercises must be performed, the number of repetitions for each exercise, session duration and its schedule (date and time). Exercises have predefined names (e.g. supination and pronation, radial and ulnar deviation) that must be provided by the application, using a select box for instance, so that the therapists do not need to insert them manually.

- Evaluate sessions: after a session is complete, a recorded video of patients performing the prescribed exercises will be sent to their therapists. Therapists must be able to use the video annotations capabilities (presented in this paper) to evaluate and send the annotated videos back to their patients. As a specific requirement of mirror therapy, therapists must be able to send videos back with the affected limb's side hidden, so patients cannot see their paretic limb and be influenced by the video.

### 5.2 Application for Patients

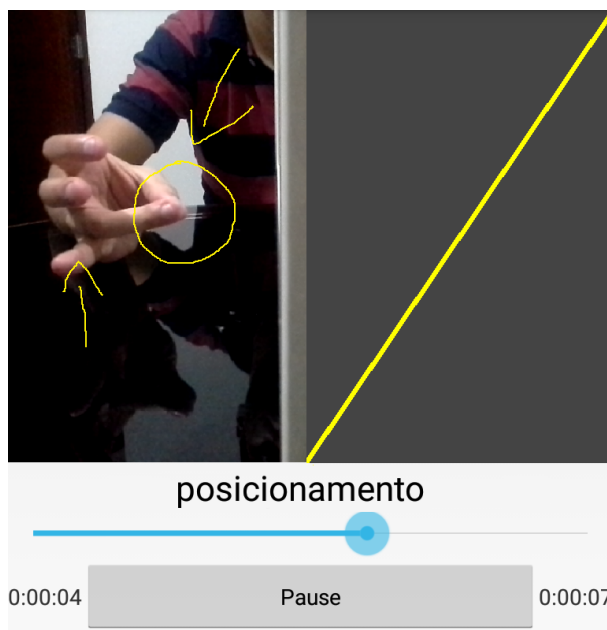
The patients' application, named Interactive Occupational Therapy, allows patients to record their exercises, send to a therapist and receive evaluated videos annotated by therapists. We describe the requirements identified with the therapists as follow.

- Perform sessions: sessions planned by therapists are sent to the their patients. Patients must to be able to visualize planned sessions and decide whether to perform or delay a session. Each session must show a planned date, a group of exercises and the number of repetitions for each exercise. When defining a session, therapists use exercises predefined names. These exercises are defined by specific names that, although well known to therapists, might not be familiar to patients in the beginning of treatment. To improve recollection, figures that illustrates the exercises are shown along with their names. Figure 2 shows an example. For recording mirror therapy exercises, the device has to be aligned with the mirror so that each side will show a recording of a different hand. Therefore, a line is shown over the camera preview for helping patients to align the device. Before starting a session, patients must be informed about its duration and receive instructions about how to place the device for recording. During recording the screen turns black. This is a specific requirement since therapists said they want to be sure that patients will not look to the screen and see the affected hand by accident. Audio instructions or counting should not be played during recording: the most important thing is that patients must be fully focused during their exercises and audio instructions might be a distraction. After recording, the video is automatically sent to the responsible therapist. Patients cannot see the video until they are evaluated by their therapists.
- Review evaluated videos: after a video is annotated by therapists, annotations have to be automatically sent to the respective patient. Patients can reproduce any video already evaluated by their therapists. Before sending a evaluation, therapists can decide whether patients can view the affected hand side. If they opt to not show it, the affected hand side is hidden on the video. Figure 3 is an example of an annotated video with a hidden side.
- Video tutorials: names and illustrations may not be enough for patients to recall how to perform exercises. The application has video tutorials for all exercises that therapists may prescribe.
- Discrimination of laterality: A set of hand images is presented for user so he must decide whether that image is a right hand or a left hand. At the end of the test, the app displays a message revealing how many images the user hit. This feature can help the patients to warm up the brain for mirror therapy.
- Personal messages: it is the same functionality presented in the therapists application. It must allows patients and their

therapists to send and receive private messages inside the application.



**Figure 2:** An example of a session with four exercises in the Interactive Occupational Therapy system. Left column: icon for each exercise. Middle column: exercise name. Right: instructions on the number of repetitions for each exercise.



**Figure 3:** A example of an annotated video that a patient can review. The video was annotated by a therapist that opted to hide the affected hand side.

To evaluate the system the applications went through usability tests and interviews about the system were conducted with occupational therapists. The purpose of these interviews was to identify the main problems and implications about the proposed remote monitoring system. In the next section we will discuss the result of this interviews.

### 5.3 Interviews with Occupational Therapists

Interviews were conducted with two occupational therapists. Being the first a PhD student in neuromotor rehabilitation area with experience in the practice of mirror therapy. The second is a professor in occupational therapy and physiotherapy with several multidisciplinary works collaborating with engineering and computing professionals. We started the interviews asking the therapists

their opinions about the performance of mirror therapy at home administered by the own patient. They answered that the greater the number of sessions the patient performs better will be the results, but it is impractical for the patient to perform the ideal number of face to face sessions per week. The training at home without the therapist is essential for the treatment. However, therapists highlighted the difficulty in identifying whether patients are actually doing the exercises at home or if they are doing the exercises correctly. The therapists also said that the performance of mirror therapy at home for long periods without any contact with therapist has a tendency to discourage patients and, therefore, increases the number of dropouts.

Then we asked if the results obtained using the system could be equivalent to those obtained during sessions in the therapist's office. Therapists said that, despite the many features present in the system, the face to face sessions has the advantage of therapist being able to change that session in real time. They also pointed out the impossibility of performing exercises with sensory and tactile stimuli at home, since the ideal is that therapists create these stimuli. When asked about the potential problems that the system could cause to mechanisms of the therapy, the therapists said that the fact of the therapist has the option to choose which video the patient will have access or if the patient will be allowed to see the damaged limb in the video initially prevents the system of causing any damage to therapy or to the illusion generated by the mirror.

We asked if the system could help in long-term adherence of patients being treated by mirror therapy. For the therapists the system can really motivate patients who need to perform long-term therapy due the constant feedback and faster response regarding treatment. They said also that, with frequent communication and with access of patients home sessions, the system could help to reduce the number of displacements of the patients and, in some cases, extend the time between the returns to the therapist's office.

To finish the interview, we asked if the system would be more suitable for any type of patient, considering aspects like age, level of education, social class or disability. For them, the only limitation would be patients access to mobile devices with sufficient resources to operate the system, or access to Internet with good download and upload speed. Therefore, social class could be a limiting factor. However, therapists emphasized that the system may be indicated for patients of all ages and varying degrees of education or performing various types of treatment by mirror therapy.

Interviews with therapists allowed an initial analysis of the results of system usage. Other studies, including testing with patients on treatment, will be performed to determine the true efficacy of the system.

### 5.4 Discussion

The integration of video recording and annotation enhance the monitoring system, since therapists not only will know if patients performed exercises: they will also see how they were performed and will be able to give advices by using video annotations.

Regarding placing the devices for recording, therapists created cheap holders to fix the smartphones. They make holders with recycled material, mainly plastic bottles or pieces of Styrofoam.

With this experience developing a monitoring system for rehabilitation exercises we could identify very specific requirements for mirror therapy. When we consider the integration with video these requirements are even more specific, for example: hide recorded videos, hide a side of the video, help patients to align the device and avoid audio instructions. Thus, we believe that other types of therapy will generate different requirements that are specific for each treatment. Through the conducted interviews we observed

that, in the opinion of the occupational therapists, guide treatment in distance is a important benefit. However, the greatest contribution is in relation to increased patient motivation by the constant interactions and feedbacks. The trend in this type of treatment is that unmotivated patients terminate the treatment before the time planned by the therapist, thus, the system can bring benefits in the long term adhesion. Another important factor to highlight is the cost of the system compared to other computational rehabilitation solutions, having a reduced cost due smartphones and tablets are becoming more popular and affordable, is rare these days who does not have a mobile device. A mirror and a compatible mobile phone are enough to use the system without the need to purchase additional equipment.

## 6. CONCLUSIONS

In this work, we presented a mobile application for video annotation executable on Android tablets and smartphones. Currently our application is the only video annotation tool publicly available on Google Play store. We designed the application to be easy to use and adequate on different screen sizes, enabling users to capture videos, add annotations, navigate through them and share their production on smartphones or tablets. We planned to use the mobile annotation tool to capture and annotate rehabilitation exercise videos. Specialists from different fields prescribe rehabilitation exercises for their patients to recover or improve specific skills. Patients should repeat the exercises daily, but the presence of a therapist is not always possible and many patients have issues when performing their prescribed exercises alone. To understand and evaluate this scenario, we relied on a collaboration of a specialist in occupational therapy. This particular specialist already used smartphones to capture videos of her patients' exercises. With the collaboration of the occupational therapy specialist, we evaluated the proposed annotation tool and identified new issues and requirements. We highlight as the main requirement the ability of integrating monitoring features into the tool.

To explore monitoring functionalities for the rehabilitation scenario, we worked with two specialists to develop a monitoring system integrated with our annotation tool. We defined our requirements considering a particular type of rehabilitation therapy, the mirror therapy. The ability to record rehabilitation videos and send them to therapists to evaluate is a great support to distance therapy. However, we observed that many requirements depend on which therapy will be conducted. In our case, mirror therapy revealed many particularities that we did not expect. We believe that other types of therapies will demand different functionalities for monitoring integrated with video. Our requirements can serve as a first step for designers of video capture and annotation systems aimed for different types of therapies.

## 7. ACKNOWLEDGMENTS

We thank Coordination for the Improvement of Higher Education Personnel (CAPES), São Paulo Research Foundation (FAPESP, grant #2011/17040-0) and National Council for Scientific and Technological Development (CNPq) for financial support.

## 8. REFERENCES

- [1] G. D. Abowd. What next, ubicomp?: Celebrating an intellectual disappearing act. In *Proceedings of the ACM Conference on Ubiquitous Computing*, pages 31–40, 2012.
- [2] G. D. Abowd, M. Gauger, and A. Lachenmann. The family video archive: An annotation and browsing environment for home movies. In *Proceedings of the ACM SIGMM International Workshop on Multimedia Information Retrieval*, pages 1–8, 2003.
- [3] S. Azenkot, S. Prasain, A. Borning, E. Fortuna, R. E. Ladner, and J. O. Wobbrock. Enhancing independence and safety for blind and deaf-blind public transit riders. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3247–3256, 2011.
- [4] M. Bakopoulos, S. Tsekeridou, E. Giannaka, Z.-H. Tan, and R. Prasad. Mobile video annotation for enhanced rich media communication during emergency handling. In *Proceedings of the International Symposium on Applied Sciences in Biomedical and Communication Technologies*, pages 32:1–32:5, 2011.
- [5] J. C. Bollen, S. G. Dean, R. J. Siebert, T. E. Howe, and V. A. Goodwin. A systematic review of measures of self-reported adherence to unsupervised home-based rehabilitation exercise programmes, and their psychometric properties. *BMJ open*, 4(6):e005044, 2014.
- [6] D. Cabral, J. a. Valente, J. a. Silva, U. Aragão, C. Fernandes, and N. Correia. A creation-tool for contemporary dance using multimodal video annotation. In *Proceedings of the ACM International Conference on Multimedia*, pages 905–908, 2011.
- [7] L. Carriço, M. Sá, and P. Antunes. Proactive psychotherapy with handheld devices. In *Proceedings of the International Conference on Enterprise Information Systems*, pages 27–34, May 2006.
- [8] R. G. Cattelan, C. Teixeira, R. Goularte, and M. d. G. C. Pimentel. Watch-and-Comment as a paradigm toward ubiquitous interactive video editing. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 4(4):28:1–28:24, 2008.
- [9] Y.-J. Chang, S.-F. Chen, and J.-D. Huang. A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. *Research in Developmental Disabilities*, 32(6):2566–2570, 2011.
- [10] P. Concejero, J. Munuera, and M. Lorenz. The MESH mobile video annotation tool. In *Proceedings of the Nordic Conference on Human-computer Interaction*, pages 583–584, 2008.
- [11] A. Cross, M. Bayyapuredi, D. Ravindran, E. Cutrell, and W. Thies. VidWiki: Enabling the crowd to improve the legibility of online educational videos. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 1167–1175, 2014.
- [12] B. C. Cunha, D. Pedrosa, R. Goularte, and M. d. G. C. Pimentel. Video annotation and navigation on mobile devices. In *Proceedings of the 18th Brazilian Symposium on Multimedia and the Web*, pages 261–264, 2012.
- [13] B. C. R. Cunha, O. J. Machado Neto, and M. d. G. C. Pimentel. A heuristic evaluation of a mobile annotation tool. In *Proceedings of the 19th Brazilian Symposium on Multimedia and the Web*, pages 89–92, 2013.
- [14] B. C. R. Cunha, O. J. Machado Neto, and M. d. G. C. Pimentel. MoViA: A mobile video annotation tool. In *Proceedings of the 2013 ACM Symposium on Document Engineering*, pages 219–222, 2013.
- [15] A. C. de Barros, J. a. Cevada, A. Bayés, S. Alcaine, and B. Mestre. User-centred design of a mobile self-management solution for Parkinson's disease. In *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia*, pages 23:1–23:10, 2013.



- [16] C. Dohle, J. Püllen, A. Nakaten, J. Küst, C. Rietz, and H. Karbe. Mirror therapy promotes recovery from severe hemiparesis: A randomized controlled trial. *Neurorehabilitation and Neural Repair*, 23(3):209–217, 2009.
- [17] E. Frøkjær and K. Hornbæk. Cooperative usability testing: Complementing usability tests with user-supported interpretation sessions. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, pages 1383–1386, 2005.
- [18] T. Georgiev. Investigation of the user's text reading speed on mobile devices. In *Proceedings of the International Conference on Computer Systems and Technologies*, pages 329–336, 2012.
- [19] R. Goularte, J. A. Camacho-Guerrero, V. R. Inacio., Jr, R. G. Cattelan, and M. d. G. C. Pimentel. M4Note: A multimodal tool for multimedia annotations. In *Proceedings of the Joint Conference of the Brazilian Symposium on Multimedia and the Web Latin American Web Congress*, pages 142–149, 2004.
- [20] M. K. Holden. Virtual environments for motor rehabilitation: review. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, 8(3):187–219, 2005.
- [21] T. Huang and G. Fox. Collaborative annotation of real time streams on Android-enabled devices. In *Proceedings of the International Conference on Collaboration Technologies and Systems*, pages 39–44, 2012.
- [22] J. Lahti, U. Westermann, M. Palola, J. Peltola, and E. Vildjiounaite. MobiCon: Integrated capture, annotation, and sharing of video clips with mobile phones. In *Proceedings of the Annual ACM International Conference on Multimedia*, pages 798–799, 2005.
- [23] H. Longzhu and X.-G. Yuan. Experimental research of evaluation of temperature ergonomics of EVA spacesuit glove. In *Proceedings of the International Conference on Digital Human Modeling*, pages 929–938, 2007.
- [24] P. S. Lum, C. G. Burgar, P. C. Shor, M. Majmundar, and M. V. der Loos. Robot-assisted movement training compared with conventional therapy techniques for the rehabilitation of upper-limb motor function after stroke. *Archives of Physical Medicine and Rehabilitation*, 83(7):952–959, 2002.
- [25] M. Lux and M. Riegler. Annotation of endoscopic videos on mobile devices: A bottom-up approach. In *Proceedings of the ACM Multimedia Systems Conference*, pages 141–145, 2013.
- [26] J. S. Miller, A. Litva, and M. Gabbay. Motivating patients with shoulder and back pain to self-care: can a videotape of exercise support physiotherapy? *Physiotherapy*, 95(1):29–35, 2009.
- [27] L. Oujamaa, I. Relave, J. Froger, D. Mottet, and J. Pelissier. Rehabilitation of arm function after stroke. literature review. *Annals of Physical and Rehabilitation Medicine*, 52(3):269 – 293, 2009.
- [28] M. Rastall, B. Brooks, M. Klarneta, N. Moylan, W. McCloud, and S. Tracey. An investigation into younger and older adults' memory for physiotherapy exercises. *Physiotherapy*, 85(3):122–128, 1999.
- [29] J. H. Rimmer, M.-D. Chen, J. A. McCubbin, C. Drum, and J. Peterson. Exercise intervention research on persons with disabilities: what we know and where we need to go. *American Journal of Physical Medicine & Rehabilitation / Association of Academic Physiatrists*, 89(3):249–263, 2010.
- [30] K. N. Truong and G. D. Abowd. Stupad: Integrating student notes with class lectures. In *CHI '99 Extended Abstracts on Human Factors in Computing Systems*, pages 208–209, 1999.
- [31] E. van Leer and N. P. Connor. Use of portable digital media players increases patient motivation and practice in voice therapy. *Journal of Voice*, 26(4):447–453, 2012.