SHENKAR — שנקר
ENGINEERING. DESIGN. ART . הנדסה. עיצוב. אמנות
The Pernick Faculty of Engineering . הפקולטה להנדסה ע"ש פרניק
המחלקה להנדסת תוכנה
Software Engineering Dept.

SyncAlong

Software Engineering B.Sc. Final Project

SRS - Software Requirements Specification

Authors:

Name	ID	Email	Phone
Shani Levi	302853619	shanilevi01011@gmail.com	0543247884
Ayman Wahbani	209138155	aymanw199816@hotmail.com	0508241000

Supervisors:

Dr. Marcelo Sihman & Mr. Eyal Nussbaum

Date: 13/12/2021

Approved by: Eyal Nussbaum Signature:

Approved by: Marcelo Sihman Signature: Marcelo Sihman

Table of Contents

1. Introduction

Revolutionizing physical activity for elderlies through side-by-side video analysis.

1.1. Overview

Synchronization is widely observed in human communication and is considered to be important in generating empathy during face-to-face communication. However, the quantitative relationship between body motion synchronization and degree of empathy is not fully understood.

Synchronization of non-verbal behaviors has psychologically positive effects. For instance, body synchronization between a counselor and a client represents their mutual empathy and relates to their level of satisfaction with counseling. In a group activity, when a group of students is rhythmically synchronized, they feel rapport with their group mates, a sense of belonging to the group, and a strong sense of unity. Body movement imitation between a teacher and students in educational settings leads to higher levels of rapport and greater satisfaction with learning outcomes.

Clarifying the relationship between changes in physical indicators and degree of empathy would improve interpretation of the cognitive relationship between body motion synchronization and degree of empathy from a physical aspect. Because body motions give different impressions, depending on speed and generation timing (Description given by Mehrabian and Williams, 1969; Miller et al., 1976).

Although the benefits of increased physical activity and exercise are universally recognized, many older persons remain sedentary, and relatively few achieve recommended levels of activity. Effective interventions to reverse the lack of physical activity in older adults are clearly needed.

Between synchronized activity and mental state together with the psychological benefits of synchronized activity there is a real need for a synchronization measurement system because studies show that when the degree of synchronization is high then the degree of interpersonal empathy is high, investigating body movement, phase differences, frequency of synchronization, and body movements in relation to the degree of empathy.

1.2. Problem Description and Motivation

Exercise and physical activity are thought to be among the most important lifestyle factors for the maintenance of health and prevention of premature disease and mortality. Yet sedentary lifestyles are common. Many people avoid exercise, and have done so across their lifespan. Exercise and physical activity are also considered important for positive psychological functioning.

Key facts about Physical activity

- Physical activity has significant health benefits for hearts, bodies and minds
- Physical activity reduces symptoms of depression and anxiety
- Physical activity enhances thinking, learning, and judgment skills
- Physical activity improves overall well-being

Although physical activity is a noninvasive and cost-effective method of improving the quality of health, global statistics show that only a few middle-aged and older adults engage in the recommended physical activity. This is **due to a lack of motivation and companionship.**

Synchronization in joint physical activity has psychologically positive effects such as:

- Creating a positive feeling
- Creating a sense of belonging
- Reducing feelings of loneliness
- Increasing motivation in joint activities
- Increasing perseverance for a common goal (physical activity)

1.2.1. Main goals

1. A platform that enables synchronized activity with loved ones remotely
2. A system for older adults for increase motivating to do more sport activities
3. A multi-participant training system designed to support interpersonal training in a synchronized manner between the participants.
4. A platform for communication and sports activities with adults who cannot physically attend joint meetings.
5. Positive response to participants when syncing in motion for the purpose of raising motivation.

1.2.2. Secondary goals

- View data on synchronized activities.
- Calculating and identifying synchronized movement and providing positive feedback to the user in real time.
- Raising motivation and improving the mental and health condition of adults.

1.3. Scope

The scope of the SyncAlong system is about synchronized joint physical activity.
Addresses a few key points:

- Real-time Communication
- Client-server computing or networking
- 3D human pose estimation involves estimating the articulated 3D joint locations of a human body.

1.4. Glossary

Terms:

Outcome	Numerical result of the synchronization calculation
Synchronization of movement	Synchronization of movement is defined as similar movements between two or more people who are temporally aligned.
Sync Rang	Defined range for synchronization indices in the range 60-100 and within it sub-settings for synchronization level
synchronization (High sync)	Very close approximation of Synchronization
Low sync	Not nearly the sync setting
Non-synchronization	Inconsistency of pose estimation or motion time estimation in performing shared motion between users
Suggestion (Improve Result)	A drawing showing a correction on top of a video taken
Period time	A period of time that elapsed from the moment an body part is in a certain place until the moment it returns to the position
Feedback	Numeric / visual / verbal view of the quality of synchronization between users
section	A list of varying sizes (few parts) of workouts whose total predefined time
coach	A system function which builds an automatic training set for the user and gives feedback at the end of the section

Synonyms

Outcome	Score
Participant	User
Cloud	Remote server with processing power
Session	Sportive activity, physical activity
Coach	Trainer

2. General Description

SyncAlong provides support for synchronized joint physical activity. Key elements of synchronized movement are: The similarities between the users and the temporary alignment of movements.

2.1. User Characteristics

2.1.1. Stakeholders: Client Description

- Kadar Center
- Elderly, family, and friends. The system is designed primarily for the elderly persons who are alone at home with no physical or mental possibility of physical activity or any joint activity.

2.1.2. End-Users Description and Scenarios

User Scenario 1: Jasmine, 26, lives far away and her beloved grandmother, Roberta. Jasmine is interested in performing a joint sporting activity with Roberta for the purpose of a joint positive activity and for the purpose of maintaining a healthy lifespan for Roberta. Jasmine opens new users in the SyncAlong system for herself and her grandmother and defines their data for the purpose of adapting the system to sports activity offers tailored for them.

User Scenario 2: Roberta is 80 years old, lives alone and suffers from loneliness and unwillingness to perform sports activities that are necessary to strengthen her joints, Roberta uses the system for a regular meeting with her granddaughter, Jasmine, she opens a meeting and with a click connects to the regular meeting with Jasmine, the meeting begins with a training list tailored to both Roberta's and Jasmine's data and abilities.
Jasmine and Roberta are shown side by side. The system displays a simulation of the physical activity of the session, at the end of the simulation - the session begins and the session time appears on the screen. During the meeting, Jasmine and Roberta perform a synchronized hand action and a variable animation rotates between them, When out of sync and the animation disappears

User Scenario 3: At the end of a physical activity session with Roberta, Jasmine checks the results of the session. The system displays for Jasmine the list of activities she has performed in order of date and time. When entering the meeting control the system displays the shared video of the participants and allows Jasmine to go through the points in the video where unsync, low-sync and synchronization were detected. The video is displayed with the participants' posture assessment marked And presents indices of improvement or non-improvement of synchronization compared to a previous encounter that exists in Jasmine's system (*Jasmine can choose that the synchronization improvement meter will be in front of meetings with Grandma only.)

User Scenario 4(Optional): Jeremy, 52, a financial manager, spends most of his day in a sitting position, and needs to incorporate exercise during the day to maintain a healthy lifestyle. He's determined to do a lunch break exercise. However, he has difficulty maintaining motivation in activity alone. Jeremy uses

the system for shared activities and searches for connected participants who are interested in performing an activity. He chooses to enter the room with Sarah. The sports activity begins

2.2. The Approach

The SyncAlong system will evaluate synchronized activities performed simultaneously via video. Our approach will use movement detection and image analysis and will be displayed using a web system with peer-to-peer secure video.

For the video connection for open an activity session we'll use

1. **WebRTC** that leverages **RTCPeerConnection** to send stream data between browsers also known as peers.
2. **Socket.io** to be able to send control messages in rooms

This architecture is known as **JavaScript Session Establishment Protocol**. Once this signaling process is successful, data can be streamed directly peer-to-peer between the caller and callee, in other words, between the users.

/*See architecture diagram for the video: Flow 5.1*

The session will be defined with a list of physical activities tailored to the limitations and abilities of the participants. The system will identify users estimated in real-time using **pose estimation model (computer vision)** and will identify synchronized movement of users during a joint physical activity. The system will also detect hands **using the Hands estimation model** in order to support basic actions during the physical activities through hand marking ,e.g. stopping the activity using stop-hand marking.

The sync-level will be set by approximating the pose estimation and the movement time estimate of the action measured in a joint activity between users. When synchronization is detected among the participants according to the movement performed together, positive feedback will be received to the users using **augmented reality** (AR), For example a cat that appears and walks around the user.

With respect to the feelings of the participants, after finishing a physical activity meeting, the meeting data will be saved in the database and will allow the display of data of the synchronization levels attached to the video parts according to number of frames per segment and present the error in synchronization for the purpose of improving future synchronization.

Also, the participants will be asked to fill a questionnaire and provide subjective feedback about the meeting such as to rank the level of synchronization they felt during the meeting.

Synchronization Approach

Synchronization can be described from two perspectives: at the level of system dynamics and at the level of influence among system users.

At the system level, synchronization refers to the coordination in **time** among the **states** or dynamics of the elements comprising the system.

The system considers the **differences in poses**, and the consideration of the **temporal alignments** of users' movements. which is a critical component that creates visual aesthetics seen in synchronized movement.

For minimal time to analyze synchronization among users the sync-algorithm will perform each X-sec\X-frames in the live stream.

Vision-based Pose Similarity Estimation **/*See reference description: Flow 5.2*/**

Although pose similarity estimation is a critical component in the SyncAlong system, most prior work relies on the summation of the **Euclidean distances of body parts**. Humans can have different heights and body shapes, they might be in different parts in the picture: one person may have been standing close to the camera, another might have been faraway. All these problems must be solved in order to output a correct result.

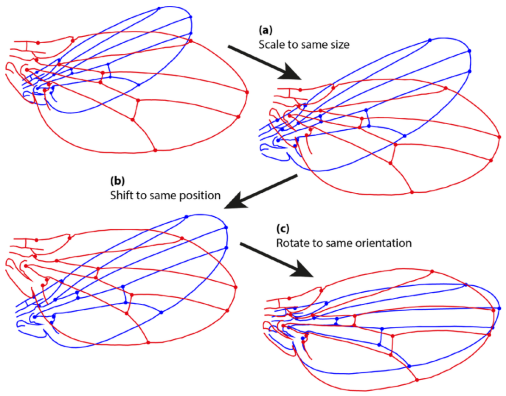
Our Approach for Human Pose Comparison

Option 1: This can be treated as a Procrustes Problem (ref-9.8)

The aim is to obtain a similar placement and size, by minimizing a measure of shape difference called the Procrustes distance between the objects.

1. Definition of 2 sets of corresponding points
   - User 1 as the model (which needs to be mimicked)
   - User 2 as the input (needs to be checked on matching grade)
2. Procrustes analysis

   steps:

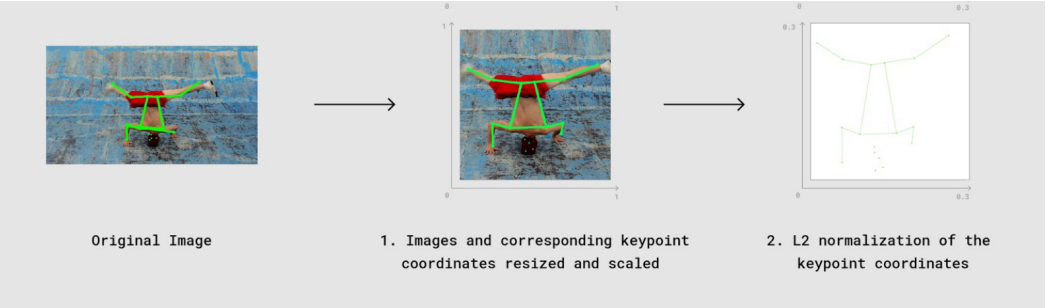


3. Using least-squares algorithm (which minimizes the sum of the squares) in order to find the affine transformation matrix

Option 2: is to take all the steps followed by Google(ref-9.3) with settings of:

Pose Similarity steps to be able to compare the data consistently:

1. **Resize and scale: we will use each person's bounding box coordinates to crop and scale each image**
2. **Normalization of the points by L2 normalization**



3. **Cosine similarity to get a similarity measure between sets of normalized key points attays.**
4. **Similarity of two vectors by Cosine and Euclidean Distance**
5. **Similarity of two time series of fram videos -** Temporal Alignment analysis

Temporal Alignment

Pose similarity considers the degree of synchronization at a video frame level. Another aspect that contributes to synchronization is the alignment of the movement timing.
Our Approach in time series analysis is to use dynamic time warping (DTW) which is one of the algorithms for measuring similarity between two temporal sequences, which may vary in speed. For instance, similarities in walking could be detected using **DTW** (ref-9.4).

Our Guiding Steps:

- Initialization
- Background subtraction (Subject Extraction)
- Motion Tracking
- Pose Estimation
- Pose Definition
- Feedback Visualization
- Performance
    - Initialization: Refers to environmental initialization, such as camera calibration, light levels analysis, environment / background.
    - Tracking: Following the movements of an object or subject of interest.
    - Pose Estimation: the estimation of the relative position and orientation of a 3D object or subject with respect to a reference camera system

2.3. Constraints

2.3.1. Time delay between different computer systems

2.3.2. Environmental configurations e.g. the placement of cameras.

2.3.3. Segmentation issues that are dependent on the potentially cluttered or obscuring background of a family room (e.g. its objects, decoration and lighting).

2.3.4. Segmentation issues dependent on the nature of the participant - adult, child, pensioner, handicapped, or other structural issues. The clothes they are wearing in relation to the background and how baggy, bulky or obscuring the clothes would affect the system.

2.3.5. Other movements in the room and their effects (e.g. non-participant occupants, pets, their shadows).

2.3.6. Changing light conditions (e.g. window light, lights turned on / off, shadow conditions).

2.3.7. The initial start-up time of the system will take a considerable amount of time for the human eye (up to a minute) for initial calculation of setting points

2.4. Assumptions and Dependencies

2.4.1. Users Settings Assumptions

a) We will consider the case with only one person on the canvas.
b) Users have a proper camera and audio connection
c) At the beginning of the joined physical activity, the camera will capture the whole body of the participant.
d) Users connected to the system via an Internet connection
e) The user's camera settings will be at 30 fps and at 360p resolution
f) When one of the participants is disconnected, the room is not closed until the two participants leave.

2.4.2. Movement Assumptions

a) The participant remains within the capture area, therefore the volume of interest is stationary and the participant is always within that volume.
b) No camera motion or constant camera motion
c) Only one person in the capture area at a time
d) The participant faces the camera at all times
e) Movement is parallel to the camera-plane
f) The participant is not occluded by the environment
g) The ground surface is flat

2.4.3. Environment Assumptions

a) The lighting is constant
b) The background is static
c) The background is uniform
d) The camera parameters are known
e) Shadow and highlight detection / removal are possible

3. Functional Requirements

Basic requirements for behavior of the SyncAlong system.

1. Users can register in the system as a trainer/instructor, regular user or a senior user.
2. The system will allow registered trainer-type users to create new users to the system and receive viewing and editing permissions of these users (add an elderly user). In this mode the system will set up a new repeating event for the two users and will provide a quick access option to this repeating session.
3. When requesting to open a room, the system will define a room for the user according to the participants data. The system will allow the user to change the room settings such as: room type, participant invitation, activity type, audio connection
4. The system will define a training set list tailored to the meeting of the participants based on their profile data and will be possible to change when requesting to open a room.
5. The system will display the physical training by presenting a short simulation before each start of a new training in the defined training set.
6. The system will maintain a list of predefined types of training according to the type of training and the type of treatment in the body part.
7. The system will display participants side by side in a web system during a joined physical activity.
8. The system will detect the pose of the body and hands of each participant to adjust the pose points for the purpose of starting point ratio.
9. During a meeting with participants and a list of training set attached to the meeting - the system will allow participants by identifying predefined hand markings to mark the system to:
    - Move to the next workout on the list.
    - Stop to rest.
    - End the session.
10. The system will monitor during an active meeting the level of synchronization between the participants every X time in the frame and will provide visual feedback to the participants' screen if there is synchronization
11. The system will save the meeting data and will allow meeting observation and receiving feedback and improving synchronization

12. The system will allow viewing of the session according to the sync level segments during the video. See Scheme 6.3.

- Shades of red: no sync
  - dark red - high sync anomaly
  - light red - low sync anomaly
- Shades of blue: there is a sync
  - dark blue - high sync
  - light blue - low sync

13. The need for accuracy. Accurate reconstruction of a physical space within a virtual space is important for several applications such as: interactive visualization of remote environments or objects, virtual modification of a real scene for augmented reality tasks

14. The application will work on PC via web interface

15. Tagging errors based on analytical metrics such as locations and angles, In the initial stage we will define for ourselves synchronization error indices is:

    * Based on a comparison with the user who opened the session

        1. Out of range of the pose-dot radius.

        2. The speed of movement in the area of the frame tested is not in the range of change up to X

    * Speed range and radius size will be defined later

16. Production environment: the application will work on PC via web interface
17. **Optional*: The system will allow users to create a joint physical activity with users by inviting a specific participant (private room) or by inviting connected participants (public room)

4. Non Functional Requirements

1. Privacy: User videos are secure and not exposed to unauthorized users
2. Extensive data storage support
3. The system will process videos on the server-side
4. Performance: Maximum accuracy of time synchronization of various devices and speed of movement while streaming video.

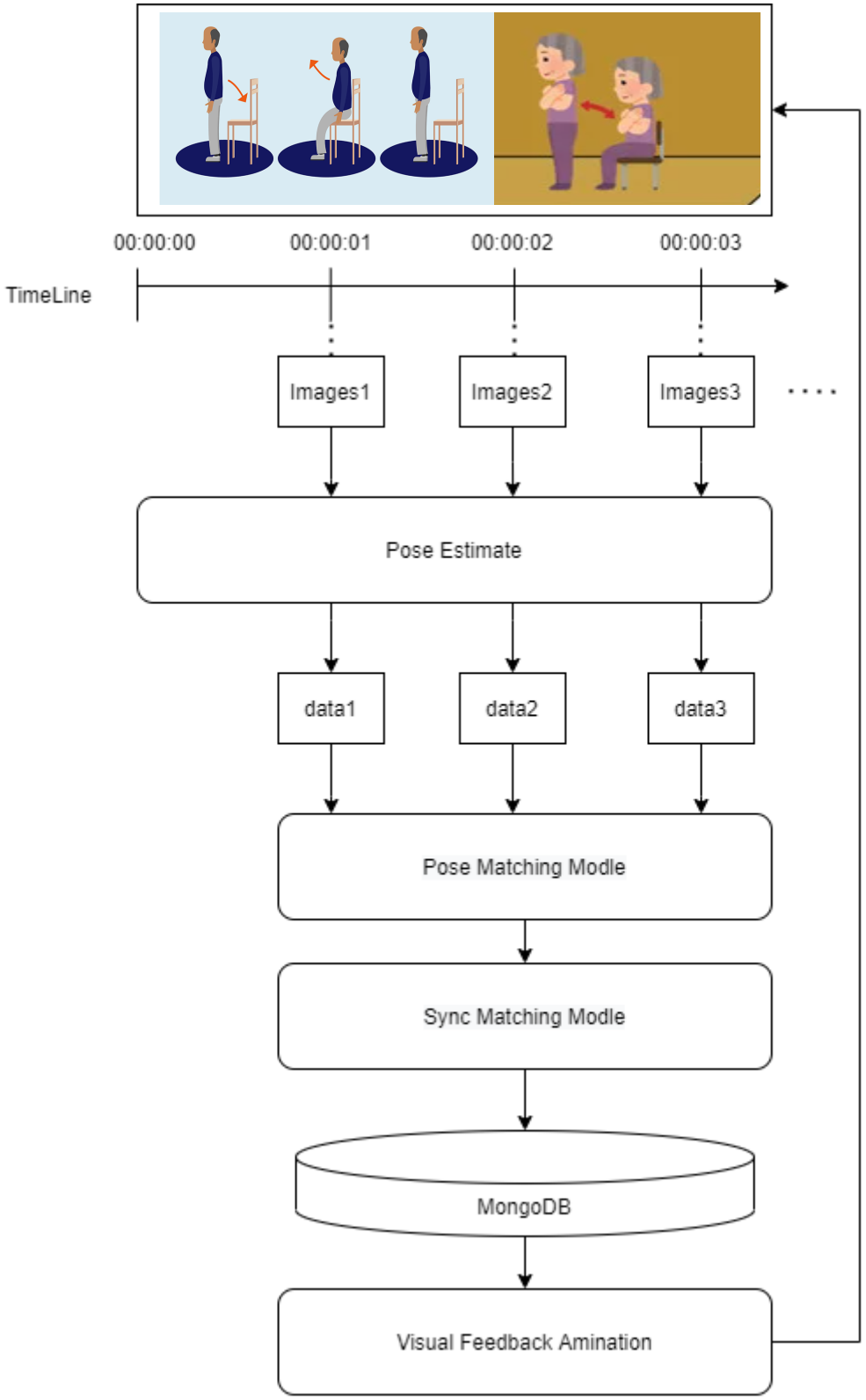
5. Resilience: Handling communication disconnection during a meeting and providing to the user who left due to communication disconnection to return to the meeting

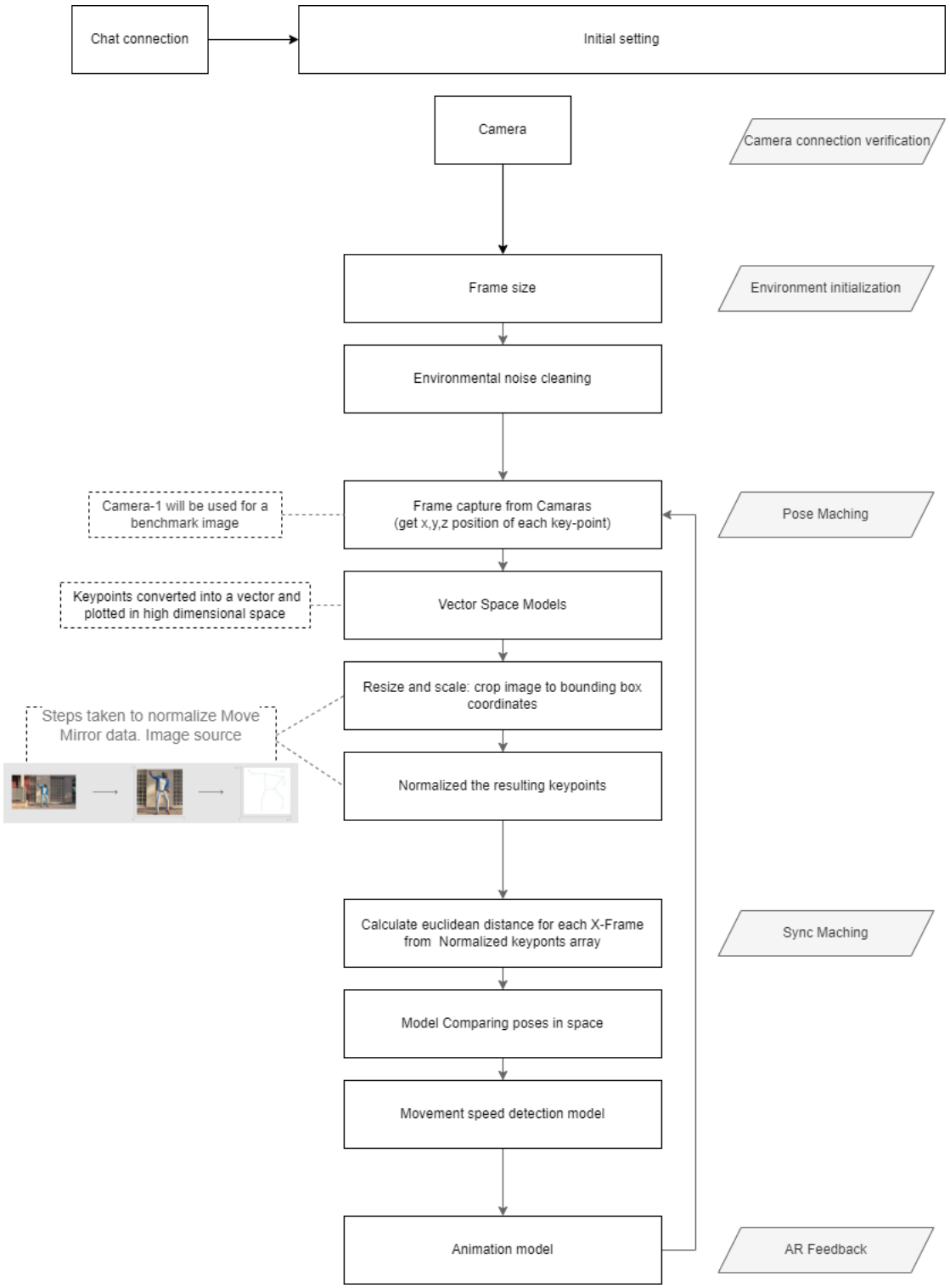
6. Data retention: Meeting data retention will be performed for all phases of the session along divided frame segments

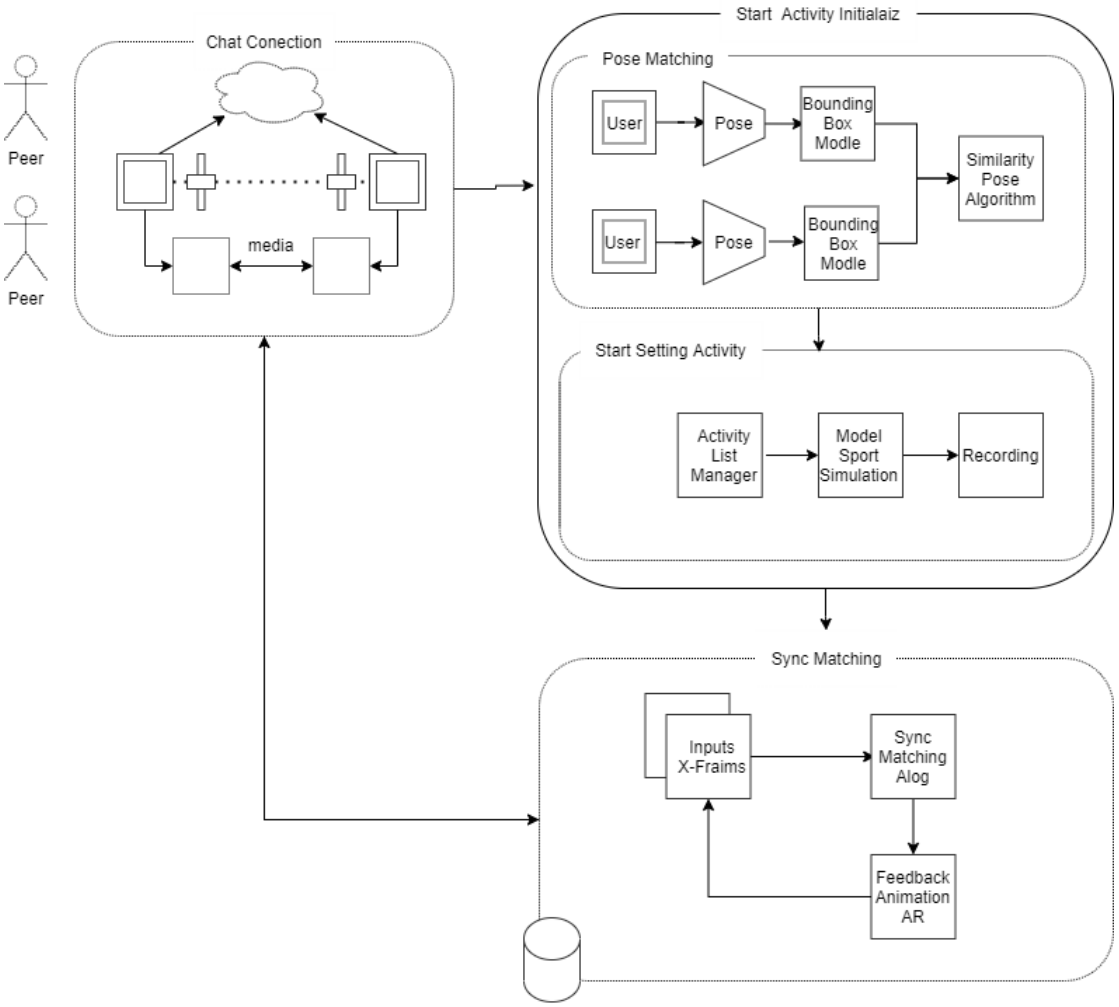
5. System Flows

Description of key system settings.

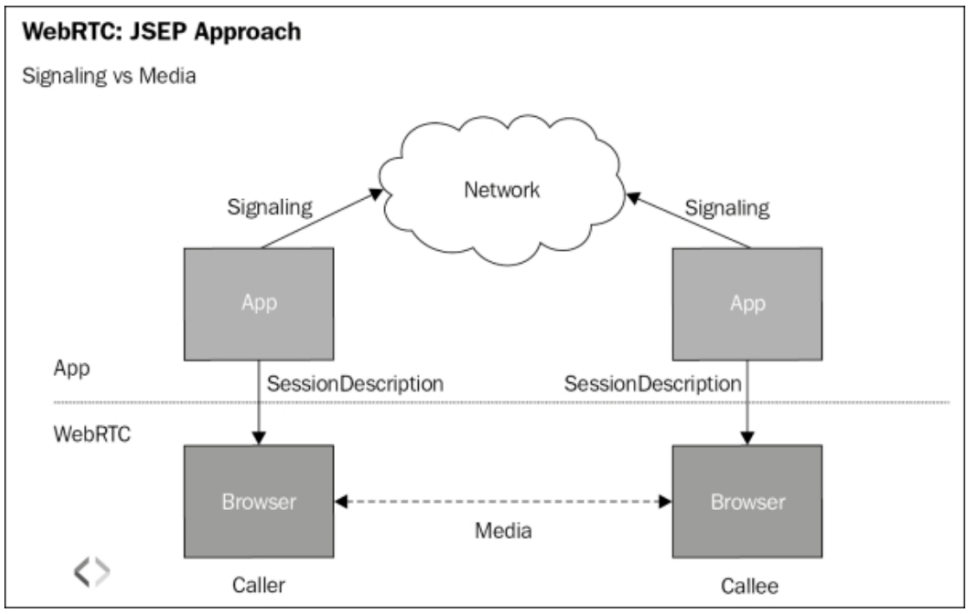
Flow 5.0 flow over all :



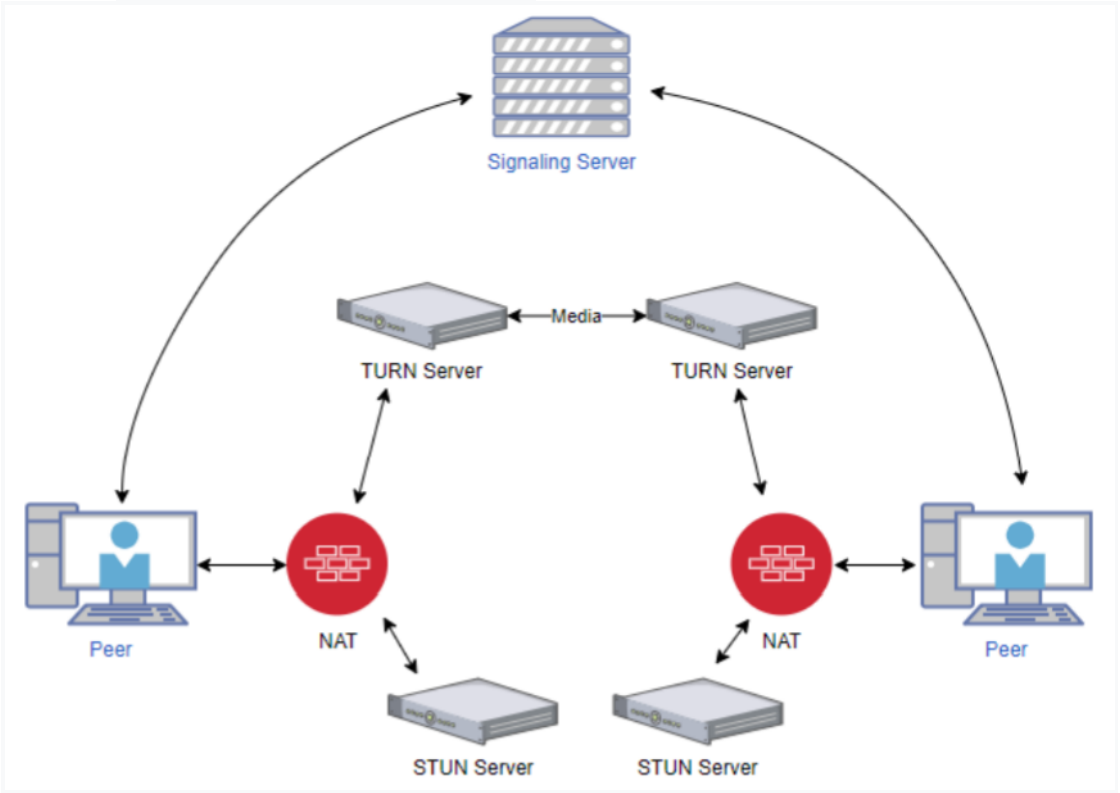

Chat connection
Initial setting
Camera
Camera connection verification
Frame size
Environment initialization
Environmental noise cleaning
Camera-1 will be used for a benchmark image
Frame capture from Camaras (get x,y,z position of each key-point)
Pose Maching
Keypoints converted into a vector and plotted in high dimensional space
Vector Space Models
Resize and scale: crop image to bounding box coordinates
Steps taken to normalize Move Mirror data. Image source
Normalized the resulting keypoints
Calculate euclidean distance for each X-Frame from Normalized keyponts array
Sync Maching
Model Comparing poses in space
Movement speed detection model
Animation model
AR Feedback


Chat Conection
Peer
Peer
media
Start Activity Initialaiz
Pose Matching
User
Pose
Bounding Box Modle
User
Pose
Bounding Box Modle
Similarity Pose Algorithm
Start Setting Activity
Activity List Manager
Model Sport Simulation
Recording
Sync Matching
Inputs X-Fraims
Sync Matching Alog
Feedback Animation AR

Flow 5.1: JSEP signaling and media , taken from website

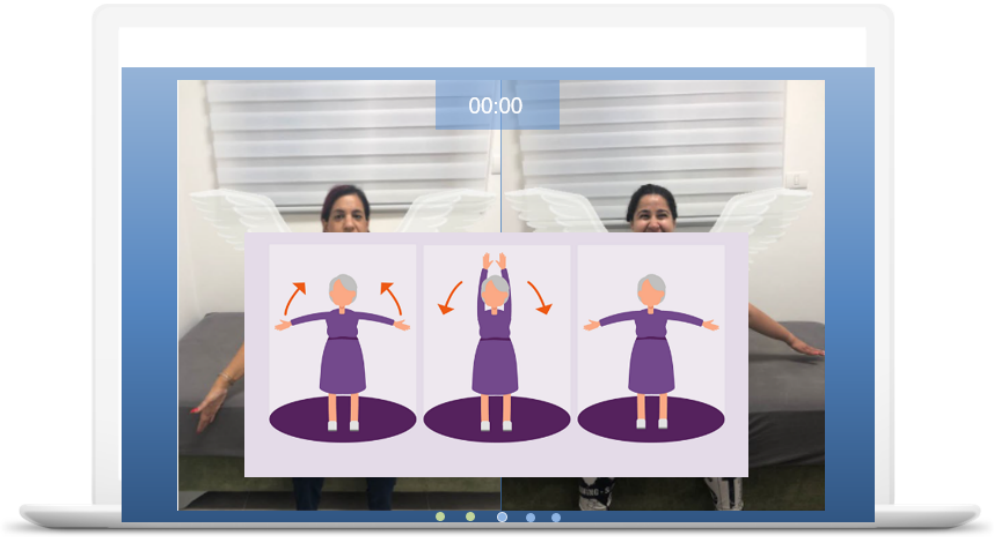


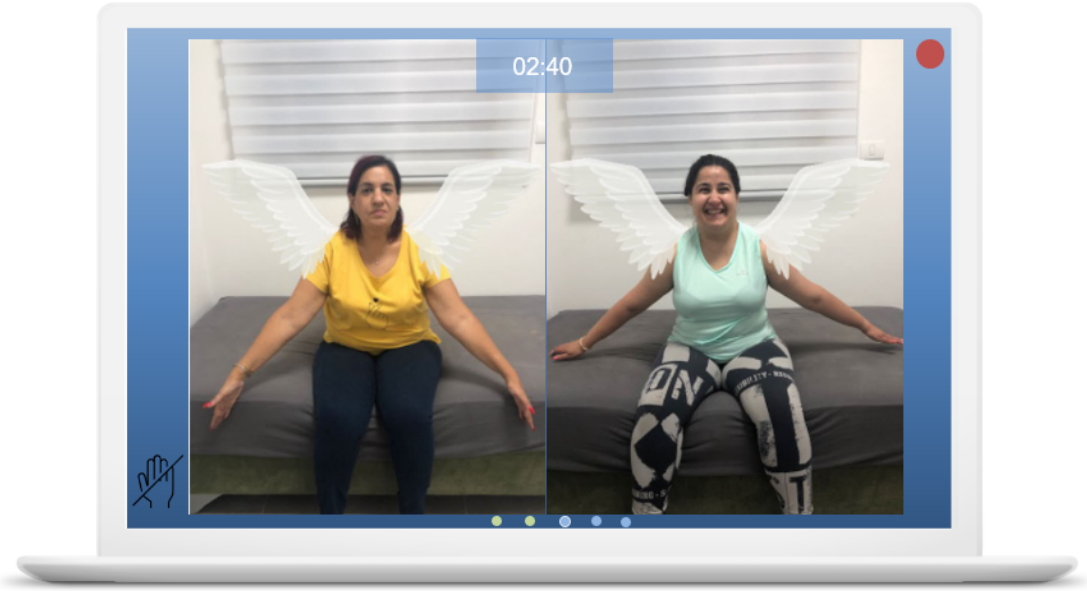
Flow 5.2: Overall per-to-per connection (link)

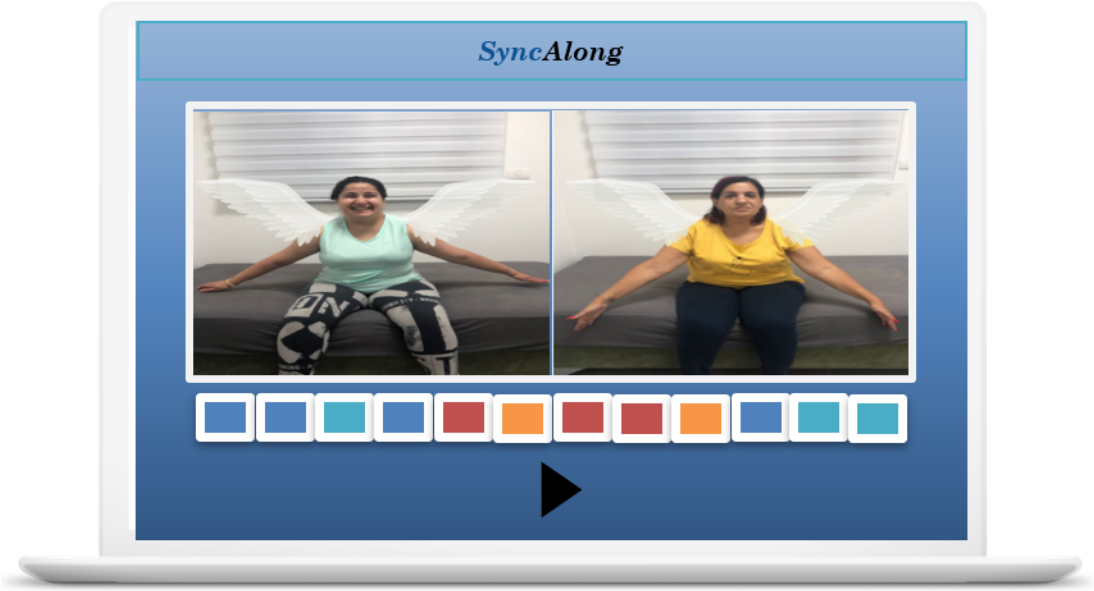


6. System Screen Specifications

© Credit to Kadar Center for the photos






SyncAlong


00:00


SyncAlong

7. Non Goals

This section of the document covers the Non goals that were not considered related to the SyncAlong System.

1. Encryption of information
2. Professional training system

8. Open Issues

1. Initializing User Dimensions: Comparison evaluation of pose estimation of participants in a joint session
2. Synchronization algorithm
3. Web network delays and network load in video streaming
4. Delay time different computer systems
5. Detection and separation of noises and movements of non-central objects.
6. Create a robust system capable of recovering from communication errors.

9. References

1	**Real-time multi-camera stream analysis**	https://www.scss.tcd.ie/publications/theses/diss/2008/TCD-SCSS-DISSERTATION-2008-021.pdf
2	**Human Pose Comparison and Action Scoring**	https://medium.com/analytics-vidhya/human-pose-comparison-and-action-scoring-using-deep-learning-opencv-python-c2bdf0ddecba
3	**Move Mirror of Google**	https://blog.google/technology/ai/move-mirror-you-move-and-80000-images-move-you/ https://medium.com/tensorflow/move-mirror-an-ai-experiment-with-pose-estimation-in-the-browser-using-tensorflow-js-2f7b769f9b23
4	**Dynamic Time Warping (DTW)**	https://en.wikipedia.org/wiki/Dynamic_time_warping https://towardsdatascience.com/dynamic-time-warping-3933f25fcdd https://github.com/MJeremy2017/machine-learning-models/blob/master/Dynamic-Time-Warping/dynamic-time-warping.ipynb https://www.npmjs.com/package/dynamic-time-warping
5	**Model_for_3D_Human_Pose_Estimation**	https://www.researchgate.net/publication/352244433_A_Synchronized_Reprojection-based_Model_for_3D_Human_Pose_Estimation
6	**Real-time DistributedChatting Application**	https://www.academia.edu/34252609/Build_a_Real-time_Distributed_Chatting_Application
7	**Chatting**	https://www.ged42.com/insights/coe/javascript/developing-real-time-secure-chat-application-whatsapp-signal-end-end
8	**Match Pose**	https://d1wqtxts1xzle7.cloudfront.net/61099639/match-pose-a-system-for-comparing-poses
9	**Human Pose Matching on mobile**	https://becominghuman.ai/single-pose-comparison-a-fun-application-using-human-pose-estimation-part-2-4fd16a8bf0d3
10	**A WebRTC based Efficient and Scalable P2P Live Streaming Platform**	https://arxiv.org/pdf/2105.07558.pdf
11	**AR**	https://medium.com/@priyaanka.garg/comparison-of-human-poses-with-posenet-e9ffc36b7427 https://blog.christianperone.com/2013/09/machine-learning-cosine-similarity-for-vector-space-models-part-iii/ https://www.partech.nl/nl/publicaties/2020/07/augmented-reality-part-2-how-to-get-started-with-ar-as-a-javascript-developer#