**Camera Evaluation**

The [Kaiser Baas x90](https://www.jbhifi.com.au/cameras/video-cameras/kaiser-baas/kaiser-baas-x90-action-camera/984438/) is the chosen action camera for our project that we have now acquired and performed initial testing with. It uses a micro SD card for storage and records with a 170º field of view allowing for the majority of actions performed in front of the wearer’s torso to be detected.

We performed an initial test of the camera, capturing footage found at the YouTube links bellow. These initial tests were conducted by following a plan to perform a variety of actions in the same order each time. The primary difference between the following 4 tests is the positioning of the camera with the chest mount due to a mixture of attempting to find the best position on the torso and dealing with camera angling issues (discussed below).

In the *‘Normal’* video the chest mount was worn in its default setup with the camera plate resting near the bottom of the sternum on the wearer. However it is also important to keep in mind that when dealing with the chest mount, the positioning is subjective depending on the wearer’s body type and gender.

* [x90 chest mount (Normal)](https://www.youtube.com/watch?v=vdvz4TughNk)

In the *‘High’* video we positioned the camera around the bottom of the neck near the top of the sternum. Initially this was to try and compensate for the angling issues but in the end turned out to provide us with what believe is the best footage from this test session despite being a bit angled downwards which we can fix in the future (discussed below).

* [x90 chest mount (High)](https://www.youtube.com/watch?v=k08XKh_1aqo)

In the *‘Stuck on’* videos we positioned the chest mount in the *‘Normal’* position but instead of the attachment piece, used a supplied adhesive piece to simply stick the camera to the plastic mount to circumvent the camera angle issues that were preventing flat, 90º recording. At the time of the test records, we believed this positioning was the best one although, upon reviewing the footage, have since determined that in order to get the widest field of view possible the *‘High’* positioning was preferred. The *‘High’* positioning was also against a flatter, more solid part of the body meaning that the camera angle would be the most consistent with different body types and genders.

* [x90 chest mount (Stuck on)](https://www.youtube.com/watch?v=jc73DVkO35s)
* [x90 chest mount (Stuck on) #2](https://www.youtube.com/watch?v=VNwryscJBII)

We also performed some basic boundary and distance/depth tests that give an idea of the limits of the camera, although we would like to re-do these in the correct position/set-up, and in more detail to provide us with better data to help inform our filming for final data set.

* [x90 chest mount (Stuck on) - Boundary Test](https://www.youtube.com/watch?v=F34jB1zRcDg)
* [x90 chest mount (Stuck on) - Distance Test](https://www.youtube.com/watch?v=bIphWpf16RI)

Upon completing the initial testing session, reviewing the acquired footage and some discussion, we have come to some conclusions on the capabilities of this camera and our applied methods thus far.

We are securing the camera to the user with a [Kaiser Baas Chest Mount](https://www.jbhifi.com.au/cameras/camera-accessories/kaiser-baas/kaiser-baas-chest-mount-for-x80-action-camera-and-gopro/554341/). The chest mount is an elastic harness that is reasonably comfortable to wear, although it can be a bit awkward to put on initially if the wearer does not know what to do. The camera is put inside its supplied case and clips in to a plastic plate located at the front of the chest mount. However, the default supplied piece used to attach the camera’s outer case to the chest mount is not naturally capable of being angled at a flat 90º due to a puzzling design choice that restricts the rotation of the case. This can be solved via makeshift methods, however can be permanently solved by acquiring an extended attachment piece such as the one found in [this pack here](https://www.jbhifi.com.au/cameras/camera-accessories/gopro/gopro-grab-bag-of-mounts-2016-edition/316223/).

At the time of the recording, we were unaware of several camera features such as the microphone and the date/time stamp that was being printed onto the video. For future recordings the microphone will be off and the stamp will not appear on the screen. Although you won’t hear audio on the current YouTube uploads because I manually muted the sound (which took a while, so this is good to find out before we did our final data sets).

As for problems that are actually related to the camera itself, while it has great 170º horizontal field of view, the vertical field of view is still limited. The problem with the vertical field of view is that actions performed outside of the camera’s direct line of sight are not captured. This mostly encompasses actions that are performed either very high up or low down. Actions that are performed so close to the body that they are underneath the camera or behind the camera are also clearly not captured. It is apparent from the test footage, that unless you are performing tasks that require your arms to be stretched out in front of your body, you are not capturing the majority of your upper arm and shoulder movements. This can pose a big problem because it gives us less visual information to work off when trying to differentiate the actions being performed.

Some of the problems identified here could be alleviated with the use of a [Head Mount](https://www.jbhifi.com.au/cameras/camera-accessories/kaiser-baas/kaiser-baas-head-strap-mount-for-x80-action-camera-and-gopro/554340/). If our x90 camera was head mounted, then it would rotate in unison with the wearers’ neck/head. It means that more actions would get recorded because you are likely to be looking at the action you are performing. When capturing footage from a chest mount, unless you tilt your chest in the direction of the action you are performing (which is unnatural and outside the scope of our project) then you won’t be able to capture those particular actions. However, at this stage of the project we have decided to go against the idea of a head mount due to comfort and practical concerns for end users/patients.

The only way to get a wider vertical angle with the chest mount would be to rotate the camera on the side so that it is recording in portrait, but then you would lose far too much horizontal data and would miss far too many actions so this is not an alternative worth considering provided the trade-off.

To summarise this, essentially the real limitations are less to do with the camera itself, and more to do with how we are utilising it. We maintain that this was the best option for the project when considering all the factors and the scope. While there are several alternatives, none of them are particularly viable.

A 360º Camera, while it can be chest mounted (surprisingly), is far more expensive (around $400 range as a minimum) and would produce a video too large to be easily processed since it is essentially multiple frames being stitched together to create a larger rectangular image. We’re not sure how it would be feasible to put this type of footage through machine learning. It is also worth noting that even though it is 360º since its chest mounted, it will be too close to the body that most of the new angles of footage we could capture wouldn’t be all that usable anyway.

You could alternatively try using a chest mount that suspends the camera away from the body and looks towards the body to capture the footage from an outside perspective. However, this also clashes with the comfort and practical concerns for the scope of our project. Having patients wear cameras on supports held out in front of them would obstruct them even more from completing their daily activities and would not provide therapists with useful results and thus, defeating the purpose of the entire project.

Google Glass was also briefly considered as a potential alternative. While it would probably provide us the best first person data set physically possible, and is not that intrusive or uncomfortable to wear, it is unfortunately no longer being supported or sold by google as of January 2015. What was released to the public was only a prototype however, so it is very possible that google will return to the market with a new product in the foreseeable future. But the google glass prototype retailed for around $1500 USD upon release so it is probably out of the question when it comes to affordable alternatives when it is compared to the x90 on a head mount which is 15 times cheaper with most likely comparable results. Once again the trade-off would most likely not be worth it.

**Further Evaluation after filming initial 2 hours of footage for data set:**

While filming the initial data set we noticed some of the following observations/issues.

**Depth/width problems:**

When reviewing the footage we filmed, we surprised ourselves with the cameras wide angle. It was able to see more width than we were expecting. This is good as it means we made the right choice of camera for capturing a wide field of view. The small negative there is that when we thought we were filming in front of a plain white wall, the camera was actually picking up other objects like chairs and people to the sides which we though were out of sight. We were aiming for simple, single colour backgrounds but ended up with some that were a little more “busy” than originally intended.

The main and basically only problem with the camera is that when actions are performed to close to the body (ie, above or below the camera) it is missed entirely or you can make out what is happening but it is not reliable and distinct enough to train a machine on. We had already identified this, for activities such as eating and drinking. The problem this time was primarily discovered while performing actions while sitting down. We decided to film some actions while seated since usually you use cutlery and write while seated. We ended up losing a lot of footage since most of the time, when you eat or write, you have your food or book right underneath you which ends up not being captured by the camera. This issue can be circumvented by adjusting the camera specifically for when the actor is sitting vs standing. But in a real world application, we wouldn’t expect patients to be constantly adjusting and repositioning the camera whenever they need to stand or sit down. It would be unrealistic.

Perhaps a similar camera that has less thickness to the case would help but I think only a 360 º or head mount would solve the problem. The advantages and disadvantages to those two approaches were addressed above.

**Camera position inconsistencies:**

To access the camera’s ports and the side menu button, the camera has to be removed from its case. But because of the way the camera clips into the chest mount, this means the case has to be tilted down so the camera can be removed. This is also necessary if we want to check the playback or what’s actually in the frame etc, before we start recording.

This makes it rather difficult to keep the camera positioned at a consistent angle between takes. The chest mount also needs to be on tightly, otherwise the mount may tilt to one side and put the frame on a slight angle. While conducting further filming for the Data Set we can be extra careful to make sure we don’t run into these problems. But again, in a real world application we wouldn’t want to rely on the patients constantly having to check to ensure the camera’s framing is correct or that there isn’t any titling with the chest mount.

Perhaps these consistency issues won’t be a problem if the machine learning can detect the action regardless, but it’s still worth noting this issue.

**Auto lighting adjust:**

The camera has a lighting auto-adjust/focus feature which I presume most camera have. What this means is that it tends to react either very subtly or very drastically to light changes.

For example. In one room we had a switch that allowed us to control the brightness of the lights. We filmed at 3 different “light levels”. However upon reviewing the footage it is very hard to tell the difference, when in the physical room itself it is a noticeable change in brightness. Then in other situations, such as where natural light is involved or when the patient moves between one room lit with a yellow light bulb into another room with white or florescent light bulbs. In this case, there can be a dramatic shift in light consistency.

The biggest problem is probably natural light, for example in the morning where there is typically a bright light shining through a window it will “white out” a lot of the frame. I am unsure if there is a way to disable the camera auto-adjust feature, there is probably some work around though. Clearly it is counterproductive to train the machine to try and recognise actions that are obscured by over exposure and also to only film inside with no windows.

**Activity consistency required:**

We also probably made the mistake or having too much variance in the actions we recorded. We should have opted for more repetition as opposed to variance. For example, there are many ways you can open a door, we filmed for all of them. Doors with different shaped handles (or push doors with no handles), cupboard doors, fridge doors, microwave/oven doors. Then, what is the action of “opening a door”? Is it simply the act of opening it while standing still? What if you walk through the door? Some people open while walking through. When closing a door most people will close it behind them once they’ve entered the room without looking back. Do we only record door closes when standing still? It would be a rare occasion to see someone open and close doors while standing still in reality.

It is questions such as those that we should have taken more time to determine before we filmed. Since we will more than likely need to film an additional two hours of footage, we will take a much greater focus on repetitious and stray away from unique actions. The list of actions we are going to performed will probably need to be more defined. For example from “turning action” -> “unscrewing a bottle cap”. This way we can collect a large quantity of data o train the machine on instead of having a large variety of different types of activities that fall under the relatively broad categories we’ve given ourselves so far.