Translating Head Motion into Attention - Towards Processing of Student's Body-Language

Mirko Raca
CHILI Laboratory
École polytechnique fédérale de Lausanne
RLC D1 740, CH-1015 Lausanne
mirko.raca@epfl.ch

Łukasz Kidziński
CHILI Laboratory
École polytechnique fédérale de Lausanne
RLC D1 740, CH-1015 Lausanne
lukasz.kidzinski@epfl.ch

Pierre Dillenbourg
CHILI Laboratory
École polytechnique fédérale de Lausanne
RLC D1 740, CH-1015 Lausanne
pierre.dillenbourg@epfl.ch

ABSTRACT

Evidence has shown that student's attention is a crucial factor for engagement and learning gain. Although it can be accurately assessed ad-hoc by an experienced teacher, continuous contact with all students in a large class is difficult to maintain and requires training for novice practitioners. We continue our previous work on investigating unobtrusive measures of body-language in order to predict student's attention during the class, and provide teachers with a support system to help them to "scale-up" to a large class.

Our work here is focused on head-motion, by which we aim to mimic large-scale gaze tracking. By using new computer vision techniques we are able to extract head poses of all students in the video-stream from the class. After defining several measures about head motion, we checked their significance and attempted to demonstrate their value by fitting a mixture model and training support vector machines (SVM) classifiers. We show that drops in attention are reflected in a decreased intensity of head movement. We were also able to reach 61.86% correct classifications of student attention on a 3-point scale.

Keywords

computer vision, head movement, attention, classroom

1. INTRODUCTION

One of the early studies of attention in classrooms showed that only 46% of students pay attention during the class [4]. Later studies raised that estimation to a more optimistic but still insufficient 67% [20]. This means that in practice the teachers are lecturing half-empty classrooms, even if all chairs are occupied. How can we help the teachers learn to recognize which chairs are empty?

Processing of social cues comes natural in human-to-human communication, but still remains an object of much research and few technical applications. The ambiguity of the medium limits our attempts, but in the scenarios where body language becomes the dominant form of expression, we are inclined to dig further into the matter. One such scenario is the classroom. We argue that computer vision (CV) technologies, in combination with machine learning approaches give us tools to scale-up teacher's attention to every student in the classroom, regardless of the class size. This would provide the teachers with a timely opportunity to address lower attentive class areas and draw students into the lecture, encouraging teacher's reflection in action.

Behaviour of people in large groups is unpredictable to an observer in most situations. The overwhelming amount of information forces us to focus on few individuals who we deem as the representatives of the group, and mental effort and training are required to re-divide the attention equally among many subjects [7]. In case of a lecture, teachers are active participants, splitting their attention between personal actions, material presentation and orchestration of the whole process [8].

In this work we started from the success of eye-tracking in predicting focus and tried to generalize it to students' head movement in the classroom. Birmingham et al [3] illustrate the social aspect of gaze – given an image, people first analyse the gaze, then the head and finally the posture of the people in the image to collect information about where to focus their attention. Langton [13] showed that we combine the input from head and eyes into a single stimulus. These two observations together gave us the ground to consider head orientation as *i*) informative to other humans, and thus potentially also for our algorithms; *ii*) an approximation of human gaze on larger scales of motion.

In this paper we present our process for extracting head motion and pose features from videos of classroom audience, and our initial set of analysis of the features' quality. We will try to answer if there is a general connection between head motion and attention level? What are the features of head motion that we can use in predicting attention? How do these features change with attention levels? And finally, can we use these features to predict students attention levels?

2. RELATED WORK

The umbrella of affective computing [15] has been growing in the last 15 years, and expanding the domains of it's application. The emerging sub-field of Social Signal Processing (SSP) [24, 25] made a major point of emphasizing that encoding human social and cultural information might raise the performance of the machine algorithms aimed at understanding behaviour (e.g. analysing large sport gathering [6]).

In case of human attention, it is attributed with the ability to modulate or enhance the selected information source according to the state and goals of the perceiver, and that the "perceiver becomes an active seeker and processor of information, able to intelligently interact with their environment" [5] and can be highly relevant in a learning environment [14]. Roda et al [19] already tried to incorporate the attention indication as one of the inputs in human-computer interaction, but early attempts in the classroom were not formulated as a technology which can be wide-spread, due to their complexity [1].

Detecting and displaying the gaze direction, as one of the key indicators of focus of attention, was shown to be both useful in making the interaction feel more natural [23], and indicative of the material comprehension [21] in on-line environments. Lacking the possibility of capturing gaze in a real-life scenario, Ba et al [2] demonstrated that we can estimate the VFOA (visual focus of attention) in meetings successfully based on the head pose. In the similar scenario Stiefelhagen et al [22] showed that head orientation contributes 68.9% in the overall gaze direction (where is the attention directed) and achieved 88.7% accuracy at determining the focus of attention. This gives us the indication that head motion has potential as a focus indicator, but it does not come without problems. Deeper exploration of head motion depicts it as an ambiguous indicator. Heylen's overview [10] shows that head-signals are either very contextual-dependant or are complementary signal to the main information channel (usually – talking).

Our conclusion from the literature overview is that head motion has the potential as a low-resolution measurement which we can passively acquire to determine the attention level and/or direction of another person. To fully decode it we need contextual information which will be unavailable in our approach of passive/unobtrusive data collection [16]. The features we hope to find need to be positioned in the middle between measurable and context-dependant.

3. METHOD

Training and validation of our head detector/pose estimator pipeline was detailed in our previous work [17]. We will give a quick overview of the experiment setup and detection pipeline, and focus on the steps and problems we encountered in the later stages of data extraction.

3.1 Experiment design

We collected a total of 6 recorded sessions with 2 classes (demographic information shown in Table 1). Each classroom was observed with several cameras positioned above teacher's head around the blackboard area of the classroom (camera view of the classroom is shown in Figure 1). The

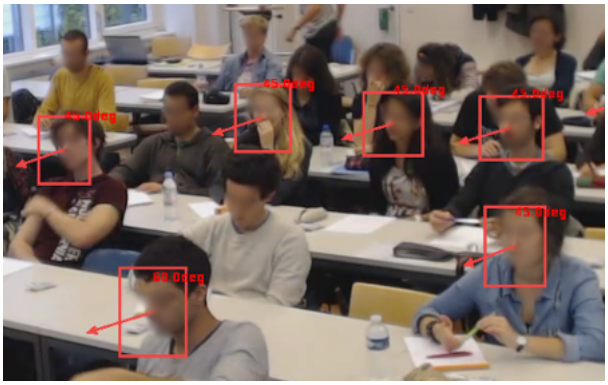


Figure 1: Examples of gaze detections, showing the classroom during the lecture.

cameras were synchronized and each student visible in the video was annotated with an unique ID (maintained over all recorded sessions) and a rectangular area of the video which the student occupies. Given that the angle of the face detected is relative to the camera viewpoint, we introduced angle offsets for each student. If a student was visible from several cameras, best quality recording was used.

Class	Size	F.ratio	Mean attend.	Sess	Cams
1	62	35.48%	39.34($\sigma = 1.15$)	3	5
2	43	34.88%	27.5($\sigma = 6.55$)	3	4

Table 1: Statistics of the two captured classes, showing the number of students, percentage of female students, attendance, number of sessions recorded and number of cameras used.

Similar to attention probing used in earlier experiments [4] we asked students to fill out the questionnaire about their attention during the class. At four different times the classes were interrupted and students recorded their attention on a Likert scale from 1–10 (details of the questionnaire design are presented in [17]). The distribution of all collected answers is shown in Figure 2. From each of the 6 processed classes we recorded 4 measurements of attention per student, associated to the time period before our interruption, duration of 7-10 minutes. In order to turn the problem into a classification one, we labelled the values of the students' responses as *low* (reported attention 1–4), *medium* (5–7) or *high attention* (8–10), based on our observations of attention distribution (regions marked in Fig.2).

3.2 Video analysis

The head-pose detection and pose estimation was built on top of the part-based model for head detection published by Zhu et al [26] which was re-trained for lower resolution images and different head poses on the AFLW dataset [12]. We trained a geometrical head-pose estimator (focusing on horizontal angle or "*pan*" of the head) by using the dlib library [11]. The precision of the estimators was checked on the Pointing'04 dataset [9]. Each detection consists of the assumed rectangle of face area, estimated angle of the face ("pan") and score (detector confidence).

The major problem for reaching the meaningful measure-

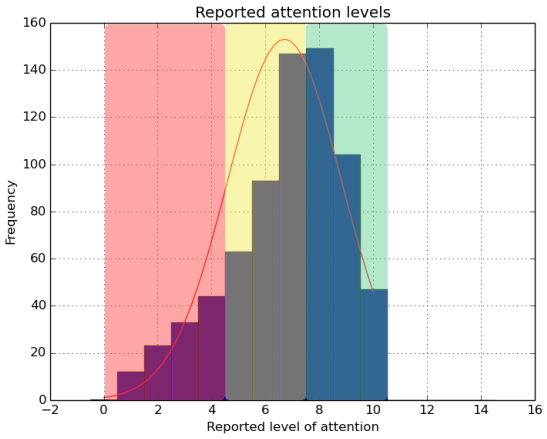


Figure 2: Histogram of all reported levels of attention with the used limits to designate the *low* (red zone, <5), *medium* (yellow 5-7) and *high* (green, 8-10) levels of attention.

ments was the instability of the detector/estimator output. The measurements were very noisy since the feature extraction step was not formulated as a tracker, which would provide temporal consistency. The second problem came from the setup itself — given the location of the cameras (around the black-board, visible in Figure 1), the subjects sit closely together. This causes a considerable amount of *i*) interpersonal occlusions and *ii*) gaps in detection and *iii*) miss-assignment of detection instances (visualized in Figure 3a).

Simple attempts to pick the best-scoring detection within the region did not yield a stable output, given that on most occasions the head of the neighbouring student would wander into the region and take over as the best detection. Fitting prior distributions (2D Gaussians) for expected head locations also did not improve the assignment, as students usually create 2 or 3 mixtures of points (depending on their sitting poses), which is indistinguishable from the case when two people occupy the given space.

Finally we settled for the formulation with labelled GMM (Gaussian Mixture Model). By taking sparsely sampled detections over time (one frame every 2 seconds) and accumulating all the detections, we depicted the overall probability of detecting faces in different positions of the camera view. The "labelled" part consists of manually specifying the relevance of each mixture in the probability, by either labelling the mixture as a specific person or miss-detection. With this we could filter-out all the irrelevant detections for a specific person by only considering detections which were assigned to one of the person-related clusters in the GMM (Figure 3b).

To improve the precision of the GMM fits, before training the model we eliminated the outlier points by thresholding the minimal number of neighbours a point needs to have in order for it to be further considered. This is possible due to the fact that the people remain in distinct positions for long periods of time, causing dense groupings of detections. The threshold was dynamically determined for each video, by eliminating the 0.5% of points with lowest number of neighbours. The major role of the GMM filtering step was to eliminate false positives, as the clusters could not always be mapped one-to-one to an individual. Additional constraints during the GMM training phase could solve this problem.

After filtering out the miss-detections, temporal consistency was ensured by using a simplified Kalman filter approach – the next detection is expected to be in the close proximity of the previous detection. If no detections were observed within a specified radius from the previous detection, the radius is increased for the next processed frame and no detection is reported, simulating the increase in uncertainty. The major differences from the Kalman filter is the absence of motion model (the face is expected to remain at the same place) and the lack of probability propagation. This enabled us to use only the real detections and not estimates, which is relevant in order to model the heads in a bow-down position. The region growing was preferred over moving Gaussian in order to put a hard limit on the detections which can be considered.

After each processed person in the video, to make sure that the detection would not be used two times, we removed the detection after it has been assigned to a person. This turns the algorithm into a greedy approach, and making the order in which the persons are processed important. We chose to process the persons from front-to-back given that each person sitting closer to the cameras is more likely to be correctly detected. After extracting detection tracks for each person, values of the detection rectangle position and gaze angle are smoothed with a "sliding window" approach.

3.3 Features extracted

The input features used in our predictions were largely based on the information extracted from the cameras, but not exclusively. All features used are shown in Table 3.3. As we noted before, the time and spatial arrangement also plays significant role in the attention estimation [18], so we included the information about the distance of the student from the teacher (distance and row fields), and time of the sample within the class (period).

We tried to model the eye contact in the class with the percentage of time that we detected the student's face in the video. Initial assumption is that this would allow us to measure the time the student spent looking down just by noting how long was the head absent. The noise in the measurement originates from the false negatives of the detector, which is dominantly influence by the distance from the camera. Even though we resorted to using zoom-lenses for the distant people in the class (which makes the measurements comparable even on the capture level to the people in the front rows), there still was a significant correlation between the row in which the student sat and percentage of time detected ($r = -0.1867$, $p = 0.009$), although it was weaker than the correlation with the Cartesian distance from the teacher ($r = -0.2137$, $p = 0.002$) which encodes width as well as depth of the classroom.

"Head travel" records the total accumulated head travel in the horizontal plane. We ignored the potential head-travel in the periods when we did not detect the face of the stu-

dent. In order to neutralize the potential influences of person's rhythm and distance from camera, we also included a normalized version of the measure, by using all the measurements of a single person to determine the mean and scaled it with the variance of those measurements. Samples with a single measurement were excluded.

We modelled the focus of the student with 3 connected measures of stillness – number of still periods, mean duration of the still period and percentage of time spent still. Stillness was defined as periods during which the head changes are less than 10°, and where the head's angle does not move away from the initial angle more than 10° (in order to prevent slow drifting to be classified as stillness). "Stillness periods" are defined as non-overlapping periods of minimum duration of 5 seconds, in which the stillness condition is true. From there we get the first two measures by counting the number of such periods and their mean duration. Percentage of time spent still is the ratio of time classified as being still over the duration of the attention period.

All measurements were considered per attention period and per person in order to associate the features to the labels acquired from the questionnaire. In case of regressions/ correlation tests, we also tested the correlation of the measures after the logit transformation, by first bounding the value scopes (finding minimum and maximum values for all measurements and scaling them to the 0.1 – 0.9 interval) and applying the $\log_e\left(\frac{p}{1-p}\right)$.

4. RESULTS AND DISCUSSION

4.1 Features

First significance tests showed the correlation between the pure attention level with the percent of time the person was detected (Pearson's $r = 0.1158$, $p = 0.01$, 577 samples). This can be explained with the idea that engaged students will maintain more contact with the activities in the classroom. Apart from being more visible, students head travel did not show significant difference on the overall scale. We expected this as the measurement itself can be easily affected by noisy measurements, even though we did take steps in smoothing the data.

Head travel became significant when testing its potential to measure the change in behaviour. After eliminating the individual differences with normalization of head travel, we found that positive changes in attention were reflected in increase in head travel (Pearson's $r = 0.21$, $p < 0.01$, 236 samples), as shown in Figure 4.

Of the measures of stillness, only "percentage of time spent still" recorded a significant, but very weak correlation (Pearson's $r = 0.09$, $p = 0.02$). After comparing it with the "percentage of time detected" we found a very high and significant correlation between the two measures ($r = 0.91$, $p < 0.01$), which does not allow for great significance of the measure. We kept the measures for further testing.

4.2 Models

Next step in demonstrating the usefulness of the features was to try to predict the attention levels based on their combinations. After initial attempts with linear regression

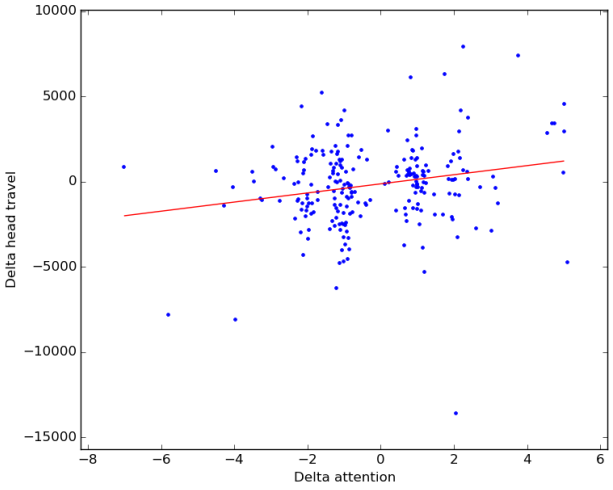


Figure 4: Change in normalized head travel correlated to the change in attention. Red line represents the linear fit. Pearson's $r = 0.21$, $p < 0.01$. Number of samples 236. Noise added for the visualization after the linear fit.

which were not successful, we switched to the mixture model. Our mixed model for *logit attention* (**A**) with *period* (**P**), *row* (**R**), *number of still periods* (**N**) and *head travel normalized* (**H**) takes form

$$L(A) = 1.061 - 0.060P - 0.128R + 0.012N - 0.035H.$$

Although its predictive power ($R^2_{random} = 0.54$ and $R^2_{fixed} = 0.05$) is limited, significance encourages further investigation of more advance supervised learning methods.

With that in mind, we tried an exhaustive search of all feature combinations and SVM parameters to achieve the best prediction of the three categories of "labelled attention" – *low* (100 samples), *medium* (270 samples), *high* (246 samples). Training of the classifiers was repeated in several rounds (500 iterations) with random drawing of training and testing samples, while making sure that the ratio of samples for each output category is maintained (roughly 16%, 44% and 40%). Our training procedure was based on the 80–20 split — 80% of the data used for training, and 20% data for testing the prediction of the trained classifier. To evaluate SVM parameters during the training we additionally split the 80% used for training into another 80–20 split. This gives us the final data configuration — 64–16–20 split, where 64% of the data was used for training, 16% for evaluating the SVM parameters during the training, 20% for the final evaluation of the trained classifier.

For each combination of features we iterated over the SVM parameters with sampling step of 0.1 (kernel type considered - *linear, polynomial, rbf*, and their relevant parameters). On the top scoring feature combinations we applied gradual refinement of the parameter sampling step (step size was reduced down in sequence 0.1, 0.01, 0.001 around the best scoring parameter values from the previous round). Four best scoring classifiers are given in Table 3, with the best result of 61.86% correct classifications (Cohen's kappa 0.30)

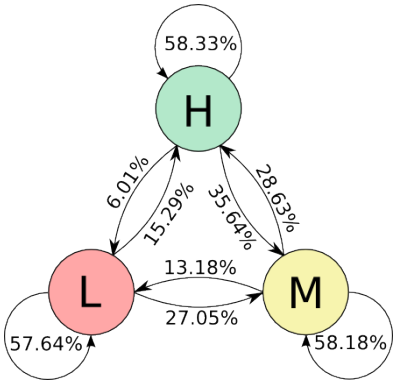


Figure 5: Transition probabilities between the three attention levels (*low, medium, high*).

on the independent test set.

Our concern was that the main informative source would rely on the *Detection percentage* or *Percentage still*, the two being highly correlated. This did happen in the early training attempts, but the features are not represented in the final set of classifiers (*Detection percentage* is used in the 10th best classifier). All of the best classifiers included a similar mix of features – head motion representatives, and some indications of distance and time of the class. *Normalized head-travel measurements* and *Mean duration of still periods* appears to be the most salient feature (both used in 3 of the 4 detectors).

Even though we saw no significant correlation of attention with class period in the feature analysis, we also tested the "attention labelled" for Markov property and got highly informative transitions probabilities shown in Figure 4.2. The trend of remaining in the same state with lower possibilities of transition to neighbouring, although not directly relevant to the attention level definitely puts additional constraints on the predictions. In order integrate this knowledge into our model, the next step was to connect our SVM predictions (observational model) and temporal consistency (transition probabilities) into a Hidden Markov Model, but due to time constraints we are unable to report the results in this publication.

5. CONCLUSION

The goal of this study was not only to answer questions about the link between student's movement and attention, but also to investigate to what extent can we approximate these variables by current techniques, without manual annotation. We defined a number of head metrics that can be extracted from a video of the audience attending a class. Considering measures that are "global" in nature (not relying on specific events such as gesturing, nodding etc.) we have shown that the change in head motion usage correlates with the change in reported level of attention. We also experimentally confirmed that higher percentage of head detection mirrors higher time spent in contact with the classroom events, indicating higher attentiveness.

For classification tasks, we found that head measurements alone were not enough to give us definitive answers about the person's attention. Each of the high-scoring classifiers used other contextual cues which related person's actions to the temporal or spacial domain (e.g. class period, distance). Also, in this report we did not explore social-level cues – how the students actions are contrasted against their immediate environment or general classroom population. We have expectations that these features will provide further contextual information, which will raise the precision of predictions.

Apart from the "global" measurements, we are also looking to explore discrete gestures which can be detected with the system (e.g. nodding, yawning, turning), of which only "bowing the head down" was used at this stage, encoded within the "percentage of time detected". The problem that we perceive is that the noise of the measurements was evident in the current setup, and that relying on the features which are more sensitive will depend on further improvements in the computer vision algorithms.

Our current conclusion is that the technology shows promise and that future investigations will bring higher accuracy and new tools to the classrooms. Our future work will try to work in parallel on finding more meaningful measures, and coordinate with the teachers to determine the best way to present the found information back to the teaching process.

6. REFERENCES

[1] I. Arroyo, D. G. Cooper, W. Burleson, B. P. Woolf, K. Muldner, and R. Christopherson. Emotion sensors go to school. In *AIED*, volume 200, pages 17–24, 2009.

[2] S. O. Ba and J.-M. Odobez. Recognizing visual focus of attention from head pose in natural meetings. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 39(1):16–33, 2009.

[3] E. Birmingham, W. F. Bischof, and A. Kingstone. Social attention and real-world scenes: The roles of action, competition and social content. *The Quarterly Journal of Experimental Psychology*, 61(7):986–998, 2008.

[4] P. Cameron and D. Giuntoli. Consciousness sampling in the college classroom or is anybody listening?. *Intellect*, 101(2343):63–4, 1972.

[5] M. M. Chun and J. M. Wolfe. Chapter nine visual attention. *Blackwell Handbook of Sensation and Perception*, pages 272–311, 2001.

[6] D. Conigliaro, F. Setti, C. Bassetti, R. Ferrario, and M. Cristani. Attento: Attention observed for automated spectator crowd analysis. In *Human Behavior Understanding*, pages 102–111. Springer, 2013.

[7] J. A. Daly and A. Suite. Classroom seating choice and teacher perceptions of students. *The Journal of Experimental Educational*, pages 64–69, 1981.

[8] P. Dillenbourg, G. Zufferey, H. Alavi, P. Jermann, S. Do-Lenhand, Q. Bonnard, S. Cuendet, and F. Kaplan. Classroom orchestration: The third circle of usability. In *International Conference on Computer Supported Collaborative Learning Proceedings*, pages 510–517. 9th International Conference on Computer Supported Collaborative Learning, 2011.

[9] N. Gourier, D. Hall, and J. L. Crowley. Estimating face orientation from robust detection of salient facial structures. In *FG Net Workshop on Visual Observation of Deictic Gestures*, pages 1–9. FGnet (IST–2000–26434) Cambridge, UK, 2004.

[10] D. Heylen. Challenges ahead: head movements and other social acts during conversations. In L. Halle, P. Wallis, S. Woods, S. Marsella, C. Pelachaud, and D. Heylen, editors, *Joint Symposium on Virtual Social Agents*, pages 45–52. The Society for the Study of Artificial Intelligence and the Simulation of Behaviour, 2005. Imported from HMI.

[11] D. E. King. Dlib-ml: A machine learning toolkit. *The Journal of Machine Learning Research*, 10:1755–1758, 2009.

[12] M. Koestinger, P. Wohlhart, P. M. Roth, and H. Bischof. Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization. In *First IEEE International Workshop on Benchmarking Facial Image Analysis Technologies*, 2011.

[13] S. R. Langton. The mutual influence of gaze and head orientation in the analysis of social attention direction. *The Quarterly Journal of Experimental Psychology: Section A*, 53(3):825–845, 2000.

[14] S. I. Lindquist and J. P. McLean. Daydreaming and its correlates in an educational environment. *Learning and Individual Differences*, 21(2):158–167, 2011.

[15] R. W. Picard. *Affective computing*. MIT press, 2000.

[16] M. Raca and P. Dillenbourg. System for assessing classroom attention. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 265–269. ACM, 2013.

[17] M. Raca and P. Dillenbourg. Holistic analysis of the classroom. In *Proceedings of the 2014 ACM workshop on Multimodal Learning Analytics Workshop and Grand Challenge*, pages 13–20. ACM, 2014.

[18] M. Raca, R. Tormey, and P. Dillenbourg. Sleepers' lag-study on motion and attention. In *Proceedins of the Fourth International Conference on Learning Analytics And Knowledge*, pages 36–43. ACM, 2014.

[19] C. Roda and J. Thomas. Attention aware systems: Theories, applications, and research agenda. *Computers in Human Behavior*, 22(4):557–587, 2006.

[20] J. R. Schoen. Use of consciousness sampling to study teaching methods. *The Journal of Educational Research*, 63(9):387–390, 1970.

[21] K. Sharma, P. Jermann, and P. Dillenbourg. "with-me-ness": A gaze-measure for students' attention in moocs. In *International Conference Of The Learning Sciences*, number epfl-conf-201918, 2014.

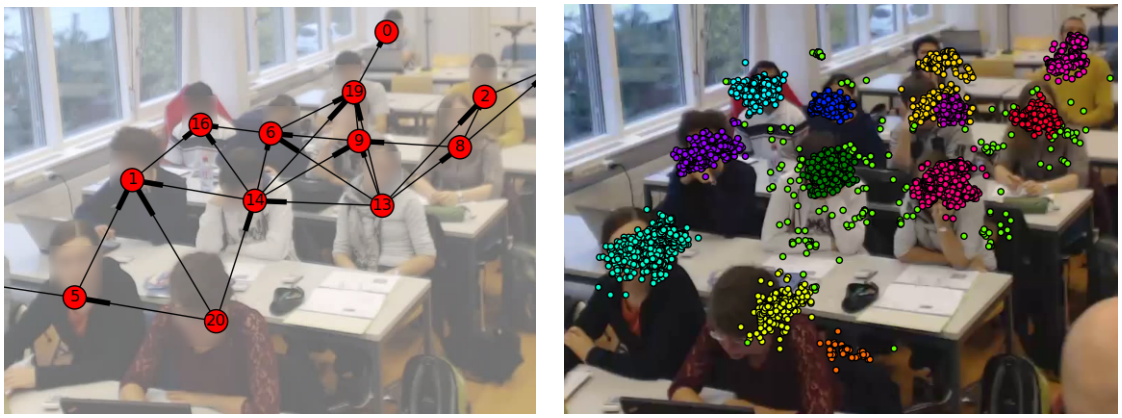
[22] R. Stiefelhagen and J. Zhu. Head orientation and gaze direction in meetings. In *CHI'02 Extended Abstracts on Human Factors in Computing Systems*, pages 858–859. ACM, 2002.

[23] R. Vertegaal. The gaze groupware system: mediating joint attention in multiparty communication and collaboration. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 294–301. ACM, 1999.

[24] A. Vinciarelli, M. Pantic, and H. Bourlard. Social signal processing: Survey of an emerging domain. *Image and Vision Computing*, 27(12):1743–1759, 2009.

[25] A. Vinciarelli, M. Pantic, D. Heylen, C. Pelachaud, I. Poggi, F. D'Errico, and M. Schröder. Bridging the gap between social animal and unsocial machine: A survey of social signal processing. *Affective Computing, IEEE Transactions on*, 3(1):69–87, 2012.

[26] X. Zhu and D. Ramanan. Face detection, pose estimation, and landmark localization in the wild. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2879–2886. IEEE, 2012.



a) b)

Figure 3: Processing of detections. *a*) Overlaps between subjects areas. Each graph edge shows neighbouring students areas and potential for miss-assignment of detections. *b*) All detections over the duration of the class, coloured depending on the cluster to which they were assigned.

Feature name	Description	Valid samples
Period	Period of the class (1–4), associated with the attention	776
Distance	Distance from the teacher on a Cartesian plane of the classroom	776
Row	Student's row in the classroom	776
Detection percentage	Percentage of the recorded time that the student was detected	668
Head travel	Accumulated changes (deltas) of the head horizontal rotations over time.	496
Head travel (norm.)	Head travel normalized over the measurements of the specific person in the class.	482
Number of still periods	Number of periods (of minimal duration of 5 seconds) during which the head movement can be considered still	668
Mean still period duration	Mean duration of the still period (as defined in the previous row)	618
Still time percentage	Percentage of time within the attention period during which the head was still.	668
Attention	Reported level of attention (1–10)	715
Attention labelled	Attention reports mapped to categories *low*, *medium*, *high*	715

Table 2: Features used in the analysis.

Kernel	Features	Score	Cohen's kappa
RBF(c=1.31, g=0.0211)	Distance, Head travel norm., Num. still periods	61.86%	0.30
RBF(c=1.21, g=0.11)	Period, Row, Head travel norm., Mean duration still	61.72%	0.32
RBF(c=1.11, g=0.061)	Head travel norm., Mean duration still	60.42%	0.28
RBF(c=1.4, g=0.04)	Period, Distance, Row, Mean duration still	59.23%	0.30

Table 3: Classifier scores for predicting "attention labelled". Score given represent the prediction score on the 20% test sample. Parameters of the kernels are abbreviated as *c* - penalty for the error term; *g* - gamma.