

AMAB: Automated measurement and analysis of body motion

Ronald Poppe · Sophie Van Der Zee ·
Dirk K. J. Heylen · Paul J. Taylor

Published online: 19 October 2013
© Psychonomic Society, Inc. 2013

Abstract Technologies that measure human nonverbal behavior have existed for some time, and their use in the analysis of social behavior has become more popular following the development of sensor technologies that record full-body movement. However, a standardized methodology to efficiently represent and analyze full-body motion is absent. In this article, we present *automated measurement and analysis of body motion* (AMAB), a methodology for examining individual and interpersonal nonverbal behavior from the output of full-body motion tracking systems. We address the recording, screening, and normalization of the data, providing methods for standardizing the data across recording condition and across subject body sizes. We then propose a series of dependent measures to operationalize common research questions in psychological research. We present practical examples from several application areas to demonstrate the efficacy of our proposed method for full-body measurements and comparisons across time, space, body parts, and subjects.

Keywords Motion capture · Human motion analysis · Measurement of body motion · Body motion analysis

Nonverbal behavior is a key ingredient in personal expression (McNeill, 1985) and the regulation of interpersonal exchanges (Ekman, 1965). Its analysis has contributed significantly to our understanding of how human interaction works. It is perhaps not surprising, then, that researchers continue to develop methods for the effective measurement and analysis of such behavior. The most common approach relies on observational coding of

Electronic supplementary material The online version of this article (doi:10.3758/s13428-013-0398-y) contains supplementary material, which is available to authorized users.

R. Poppe (✉) · D. K. J. Heylen · P. J. Taylor
Department of Human Media Interaction, University of Twente,
Enschede, The Netherlands
e-mail: r.w.poppe@utwente.nl

S. Van Der Zee · P. J. Taylor
Lancaster University, Lancaster, UK

behavior, using classification schemes that are developed to serve a particular research question (Grammer, Kruck, & Magnusson, 1998; Lausberg & Sloetjes, 2009). These schemes are often evaluative in nature, in the sense that researchers code for the occurrence of particular forms of communication, such as gestures (Doron, Beattie, & Shovelton, 2010) or facial expressions (Vick et al., 2006). Others are “physicalistic” coding procedures that utilize a more precise mapping of behavior by quantifying the movement of different limbs (Bente, 1989; Dael, Mortillaro, & Scherer, 2012; Frey & Von Cranach, 1973). While the evaluative schemes are open to issues of reliability because of the qualitative component of the coding (Scherer & Ekman, 1982), the latter physicalistic schemes have been shown to yield reliable annotations that are sufficiently detailed to animate computer characters (Bente, Petersen, Krämer, & De Ruiter, 2001). However, for both approaches, the derivation of the data through coding is time consuming, meaning that there is often an inherent trade-off between the number of coded actions and the amount of coded material.

In an effort to circumvent this difficulty, there has been a growing trend toward using technologies to evaluate behavior (Altorfer et al., 2000; Bente, Senokozlieva, Pennig, Al-Issa, & Fischer, 2008). In particular, researchers have started to undertake automatic measurement of human movement with motion capture devices. To date, such approaches have focused on examining discrete nonverbal behavior, such as head movement or gestures (e.g., Feese, Arnrich, Tröster, Meyer, & Jonas, 2012). Yet, to explore how body motion contributes to the processes of human interaction as observed in more naturalistic settings, there is a need to develop a methodology that allows for the capture over an extended period of time. In this article, we introduce a standardized approach to using motion capture methodologies for examining full-body motion. We describe how to process raw data independently of the type of motion capture device and deal with issues such as distortions, alignment, and normalization. We exemplify our approach with several case studies. As a supplement to this article, we provide MATLAB code that performs the computational steps of *automated measurement and analysis of body motion*

(AMAB), including normalization steps, distance functions, and the functionality to re-create some of the figures in this article.¹

Devices and representation

The adaptation of modern computing technology and the development of dedicated technologies has made it easier for researchers to record and analyze human body motion (e.g., Dakin, Luu, van den Doel, Inglis, & Blouin, 2010; Krishnan, Juillard, Colbry, & Panchanathan, 2009). Table 1 identifies some of the devices available for recording motion as a function of two distinctions in how they capture and treat movement data: (1) whether they rely on markers or sensors to record movement and, (2) whether they offer full-body or single-movement capture.

Marker-based technologies use a set of cameras to detect markers worn on the body. These markers are either passive, such as retroreflective balls, or active, such as infrared transmitters. The former ensure good visibility but can cause confusion across markers, while the latter use distinct frequencies to avoid confusion. For both approaches, in order to obtain a 3-D measurement of each marker, it must be visible to at least two cameras. This means that a large number of cameras are needed in order to avoid occlusion, particularly when studying social behavior.

Inertial devices overcome this drawback by measuring movements on the body—typically, through sensors worn in a suit or straps. The sensors employ changes in the magnetic field in a gyroscope-like manner to make estimates of their positions. The accuracy of this approach is typically high, although the estimated positions can suffer from drift without additional position measurements—notably, in the presence of metal in the recording environment and objects therein. Moreover, wearing a tight-fitting suit may lead subjects to be more conscious of their behavior, which threatens the ecological validity of any recorded social interaction.

An increasingly popular alternative to compensate for this validity problem is to analyze full-body movement unobtrusively using single or multiple cameras, possibly aided by projected structured light (as in Microsoft Kinect; Shotton et al., 2013). From these devices, a digital volumetric estimation of the scene and the people therein is made, in which one or more parametric body models are fitted (Poppe, 2007). The accuracy and robustness are currently lower, as compared with marker-based and inertial devices, and they suffer from the

Table 1 Overview of body motion measuring devices

Device Characteristics	Example Devices	Example Studies	
Full-Body	Type		
Yes	Marker	Vicon MX, MotionAnalysis Raptor, Advanced Realtime Tracking ARTTRACK, Optitrack Arena, PhaseSpace Impuls X2, Phoenix Technologies Inc. Visualeyex, Qualisys Oqus	Slawinski et al., 2013
Yes	Inertial	Animazoo IGS, Ascension MotionStar, Xsens MVN, YEI Technology 3-Space	Kleinsmith et al., 2011; Krishnan et al., 2009
Yes	Vision	Microsoft Kinect, Ipi Soft, Organic Motion Openstage	Mead et al., 2013
No	Inertial	Ascension TrakSTAR, Polhemus Liberty Latus, Sparkfun Electronics Witilt	Dotsch & Wigboldus, 2008; Feese et al., 2012

same occlusion problems as the marker-based approaches. However, their unobtrusive nature may make them preferable to some research designs.

The method presented in this article is generic enough to be applicable to all these techniques. Independently of the type of device used to record movement, the representation of the movements in data form is standardized. The human body is most efficiently described in terms of a series of body parts and joints, the former being shapes with a certain length and the latter being single points in space. Together, body parts and joints form a tree-like representation of the human body, and movement may be described in terms of the displacement and rotation of the joints with respect to this tree. Figure 1 shows a schematic illustration of this “kinematic tree.” The joint at the top of the tree, usually the pelvis, forms a root to which all other joints are relative. When two joints are connected to a body part, the one higher in the tree is considered the parent, and the other the child, such that joints higher in the hierarchy affect those below. For example, movement in the right shoulder affects the right elbow and wrist joints. End-effectors are joints without children (e.g., hands and head). Typically, sensors and markers are not attached at the location of the joints. While one could, in principle, use the sensor locations to analyze human movement, there is no guarantee that these locations are the same between subjects. As a consequence, motion capture equipment often employs a calibration phase to determine the joint positions relative to the sensors’ placement.

A full-body pose can be described by the rotations or positions of the joints, of which the latter is computationally

¹ MATLAB code is provided as [supplementary material](#) to this publication.

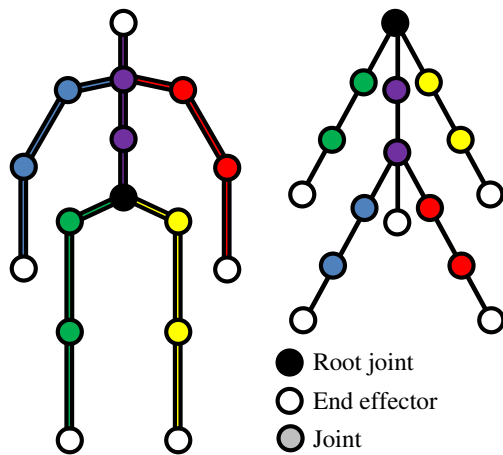


Fig. 1 Human body representation (left) and kinematic tree (right). Best viewed in color

more straightforward. Although there are a few available approaches to expressing the joint position, the most convenient for full-body capture is to use global representations, largely because they make the comparison of joint positions in time and space across subjects straightforward. In this approach, joint positions are expressed by three position values corresponding to the distance from the origin [i.e., the point (0, 0, 0)] along each of three predefined orthogonal axes (i.e., x , y , z). All devices in Table 1 can output global joint positions. The software supplied with these devices output textual representations in either XML or column format.

Once the movement has been recorded, it can be visualized in the same way as a recorded video. It is possible to have raters quantify the behavior in such visualizations by using both evaluative and physicalistic coding approaches. However, automatic measurement of body motion results in numerical representations of the body's position over time that enables a range of statistical analyses, arguably more sensitive and less prone to error than human coding. We consider the possibilities afforded by such an approach in the remainder of this article. To facilitate our discussion, we denote the k th measurement of body pose as a vector: $x^k = (x_1^k, \dots, x_n^k)$, with $k \in \{1, \dots, m\}$ for a recording with m measurements. Each component x_i^k ($i \in \{1, \dots, n\}$) of the vector corresponds to a joint position measurement along an axis. Without loss of generality, AMAB assumes that the measurements are available as a matrix with m rows and n columns. Each row corresponds to a full-body measurement x^k . The n position measurements are in fixed order of joints and axes, with each subsequent triplet of columns corresponding to the (x , y , z) values of one joint.

Data screening

Data recorded by devices such as those listed in Table 1 can be distorted in many ways. It is therefore necessary, as it is with

all inferential statistics in psychology, to screen the data prior to analysis. This process includes removal of data distortions and normalization.

Data distortions

Data distortions are due to measurement noise and longer-term inconsistencies in the data due to equipment or transmission failure. The most common noises are incidental values, which occur as a result of tracking failure (e.g., due to missing marker detections or magnetic resonance). The software receiving sensor signals often makes a first pass at removing such errors, but we propose applying a moving median filter with a modest window size—typically, on the order of 0.25–0.5 s. Using the median instead of the mean ensures that incidental off-values are not taken into account, while the window size is a trade-off between the ability to suppress “jitter” (i.e., small inaccuracies) in the output and the level of detail that is retained in the measurements. Formally, the filtered vector $x^{*k} = (x_1^{*k}, \dots, x_n^{*k})$ is obtained with

$$x_i^{*k} = \text{median}(\{x_i^{k-\lambda}, \dots, x_i^k, \dots, x_i^{k+\lambda}\}) (k \in \{\lambda + 1, \dots, m - \lambda\}), \quad (1)$$

where λ is the number of measurements before and after the current measurement that is taken into account. Figure 2 demonstrates the difference between running average and median filter, both with a window size of seven frames (0.28 s). The increase in measurement value for the running average due to an off-value is apparent.

The nature of equipment failure depends on the type of body motion device. When the time of failure is short, the missing measurements can be interpolated from the measurements before and after the failure. Linear interpolation is typically a reasonable approximation, provided that the amount of (de)acceleration is low (Poppe, 2007).

Normalization

There are a number of common analytical problems in interaction research, and these largely remain when recorded body motion is analyzed. To compare body movements within or between recording sessions, or within or between subjects, differences in body size and differences in the recording space and time must be taken into account. The most straightforward approach to removing such variations, adopted in AMAB, is to apply one or more forms of normalization.

Normalization in time

When multiple recordings are made simultaneously, synchronization is either handled by the recording software or established during data screening. The latter case occurs when recordings have been made on different computers or with

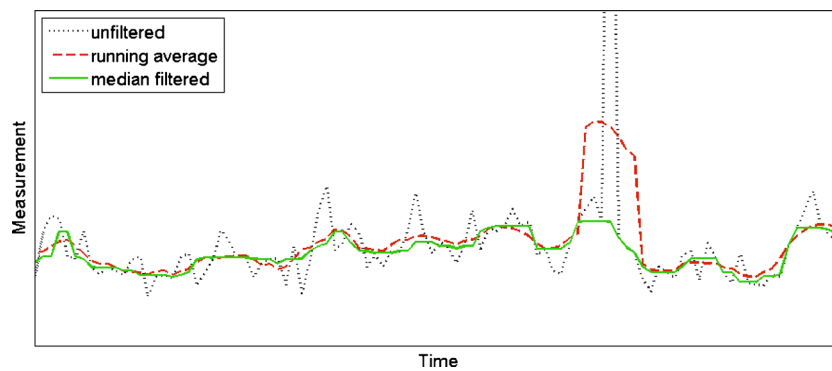


Fig. 2 Example of median filtering

different software (e.g., motion capture and video recording software). In this case, there are two possible types of normalization required: frame rate alignment and synchronization. When frame rates differ between sequences, the measurements in each sequence must be resampled equidistantly in time so that the data align to a fixed rate. The sequences may then be synchronized in time by determining the latest start point and earliest end point across the recordings and the recordings being trimmed to these points. The result is a synchronized analysis with maximum usage of the available data.

Normalization in space

The global position of a subject in the recording space affects all joint positions. This is undesirable when the body poses of a single subject are compared at different time instances or when the body postures of multiple subjects are compared. Without normalization in space, the difference in global position will influence the pairwise comparisons of the positions of each joint. Poses are normalized for position by mean centering all position measurements relative to the root of the body. Typically, the pelvis is used as the root joint P , and its location in the recording space is translated to $(0, 0, 0)$. Mean centering of the data may be applied to all other joints through the subtraction of P from the position of each joint j individually:

$$(x'_{jx}, x'_{jy}, x'_{jz}) = (x_{jx} - x_{Px}, x_{jy} - x_{Py}, x_{jz} - x_{Pz}). \quad (2)$$

Figure 3 shows an example of this kind of position normalization for the case of 2 subjects seated at opposite sides of a table. When comparing the poses of the 2 subjects, the absolute distance between them is not important. Therefore, it makes sense to apply normalization in space according to Equation 2. The result is shown in Fig. 3(I). However, the global body orientation (i.e., facing direction) of subjects typically affects comparisons between poses, which is undesirable. In our example, a researcher interested in the similarity of the subjects' poses can make an easier comparison by

rotating the pose of one of the subjects 180° around a vertical axis, as shown in Fig. 3(II). To apply this normalization in orientation, it is assumed that poses are normalized for position and that the y -axis is pointing upward. All joints are rotated around the y -axis in such a manner that the subject faces the positive x -axis. To this end, the hips are placed parallel to the z -axis. The angle of rotation θ is determined as

$$\theta = \arctan\left(\frac{x_{LHx} - x_{RHx}}{x_{LHz} - x_{RHx}}\right), \quad (3)$$

with LH and RH the indices of the left and right hip, respectively. Next, all joints are rotated around the y -axis with angle θ . For joint I with position (x_{Ix}, x_{Iy}, x_{Iz}) , the y -position (i.e., the height) remains unchanged, while the rotated x - and z -positions are determined by

$$\begin{bmatrix} x'_{Ix} \\ x'_{Iz} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x_{Ix} \\ x_{Iz} \end{bmatrix}. \quad (4)$$

An example of position and orientation normalization is shown in Fig. 4. The graph shows the sum of all pairwise joint distances (see Equation 5) between 2 subjects who approach each other, shake hands, and walk away.² The example is also included in the software that is provided with the article. Without normalization, the pose difference reflects the distance between the subjects. When walking away, one subject walks backward, while the other turns. As a result, their orientation becomes more similar, which results in a decreasing pose difference after position normalization. Finally, when poses are also normalized for orientation, the pose difference is relatively stable. Poses (I) and (II) in Fig. 4 occur after the handshake and while walking away, respectively. Although the distance between the subjects, and the difference in orientation between them, differs between (I) and (II), their body configuration is similar. This results in a similar pose difference after normalization of position and orientation.

² Subjects 18/19, trial 1, obtained from <http://mocap.cs.cmu.edu>.

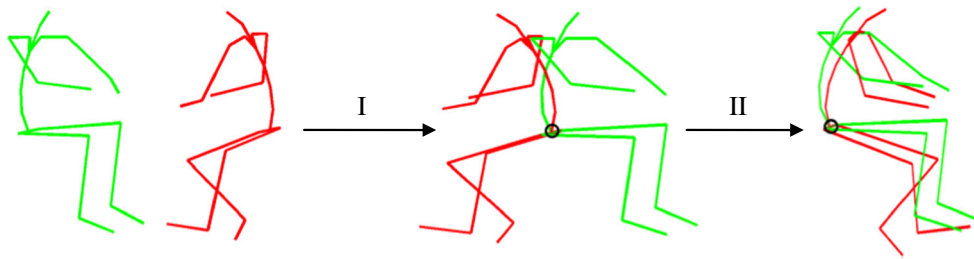


Fig. 3 Example of position normalization (I) and orientation normalization (II)

Normalization for different subjects

Subjects differ in their body sizes. Of particular concern for automated methods of measurement is limb length. These differences cause different subjects with similar joint rotations to have different joint positions and vice versa, which affects the comparison of their joint positions. To reduce this problem, the body part lengths of different subjects may be scaled to average limb sizes for a given population. Given parent and child joints P and C , respectively, and the average limb size l of the body part connecting them, the adjusted position of C can be calculated as

$$(\mathit{x}'_{Cx}, \mathit{x}'_{Cy}, \mathit{x}'_{Cz}) = \begin{pmatrix} x_{Px} + \alpha(x_{Cx} - x_{Px}), x_{Py} \\ + \alpha(x_{Cy} - x_{Py}), x_{Pz} + \alpha(x_{Cz} - x_{Pz}) \end{pmatrix}, \quad (5)$$

where $\alpha = l / \sqrt{(x_{Cx} - x_{Px})^2 + (x_{Cy} - x_{Py})^2 + (x_{Cz} - x_{Pz})^2}$, which is the ratio between the specified and the actual body part length.

Interpretation and operationalization

In this section, we demonstrate how AMAB may be applied to address various research questions that involve full-body motion measurements. We begin by describing some common variables that may be used directly by researchers as a dependent variable for their comparisons (e.g., by ANOVA), or as a basis for more dedicated measures that address specific research questions. We subsequently present four examples of comparisons across time, space, body parts, and multiple subjects, as a way of introducing the reader to what is possible through AMAB. Our coverage is by no means exhaustive, and we acknowledge that researchers must tailor their analyses beyond what we present below so that it matches their research questions. Some further examples of parameters that are similar to, and extend beyond, those presented here, are described in detail by Hirsbrunner, Frey, and Crawford (1987).

Dependent variables

For most of the research that employs full-body measurements, the operationalization of the research questions involves calculating differences between poses, movement velocity, or a quantification of the amount of body movement. Their calculation from screened data is discussed subsequently.

Pose difference

When comparing two poses A and B , their difference can be expressed as the sum of the distances between each of the joints j in the set J :

$$\delta_{A,B} = \sum_{j \in J} \sqrt{(x_{jAx} - x_{jBx})^2 + (x_{jAy} - x_{jBy})^2 + (x_{jAz} - x_{jBz})^2}. \quad (6)$$

The distance for each joint individually is calculated using the Pythagorean theorem. Pose differences can be calculated only when the sets of joints J are equal and poses have been equally normalized.

Movement velocity

The change in the pose of 1 subject is the difference between two subsequent poses and is calculated with Equation 5. Changes in pose are most conveniently expressed as velocities, with meter per second (m/s) as the unit. To this end, the pose distance needs to be calculated per second, which depends on the frame rate of the recording or the down-sampling applied in data screening. For a frame rate f and pose difference δ , the velocity v is calculated as

$$v = f\delta. \quad (7)$$

Amount of movement

The total amount of movement for a single joint or all joints can be calculated by summing all pairwise distances between subsequent measurements over an interval. For a sequence of

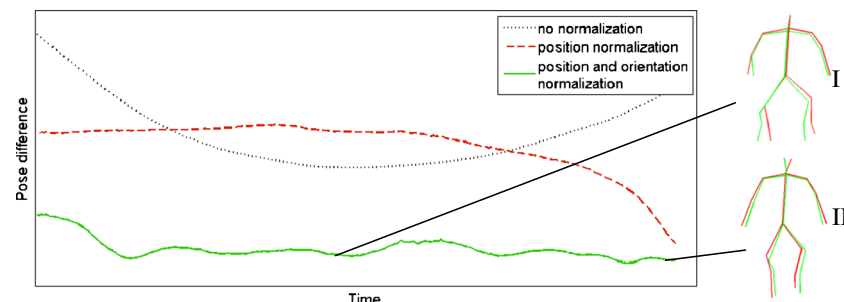


Fig. 4 Example of position and orientation normalization, with sums of pairwise joint distances shown over time

length m and all joints in J , the amount of movement α is

$$\alpha = \sum_{i=1}^{m-1} \delta_{x^i, x^{i+1}}. \quad (8)$$

When comparing the amount of movement between two sequences, these should cover time intervals of equal length and, ideally, with equal frame rates.

Comparisons across time

Motion capture devices measure full-body poses over time. While the pose is informative, one is sometimes particularly interested in the temporal aspect of human movement. Changes in pose or movement over time can be used to measure consistency of movement or response time. For example, fatigue can be measured by analyzing the decrease in total amount of movement over time. The analysis of response times to certain stimuli is particularly important to sports psychologists, who may be interested in factors that impact the swing of a bat or the speed of a start during the 100-m sprint (e.g., Slawinski et al., 2013). The latter will be used here as an example to demonstrate how the AMAB method can be applied.

Currently, response times in the start of sprinting are typically measured using pressure-sensors in the starting block (Bezodis, Salo, & Trewartha, 2010). A threshold on this pressure is set to exclude false alarms due to small changes in body pose. While pressure measurements are accurate, foot pressure is the result of all movements of body parts higher in the kinematic chain (see Fig. 1). As such, it is an indirect measurement. Ideally, one wants to analyze not only the final pressure of the foot, but also the involvement of the movement of all parts in the body. The use of motion capture devices enables one to perform such an analysis. For example, Slawinski et al. (2013) employ full-body motion capture devices to study pose change during the start of a sprint.

AMAB can be used for the numerical analysis of response times, in order to detect false starts. First, the full-body movement needs to be filtered and aligned in time relative to the starting stimulus (e.g., a gunshot). After normalization in space (Equations 1–3), frame-to-frame differences between

poses can be calculated using Equation 5. Subsequently, a threshold can be set on the movement velocity (Equation 6) in order to detect the movement offset and to prevent false alarms.

Comparisons across space

While normalization of position is typically carried out to compare different poses, the absolute position itself can also be used as an independent measurement. One could measure the position relative to an object with known location or the distance between two people. For the latter, Hall (1966) distinguished four different zones of interpersonal distance: intimate (<1.5 feet), personal (1.5–4 feet), social (4–12 feet), and public (12–25 feet). Researchers have studied interpersonal distance at all four levels, usually using video recordings to estimate distances (Hayduk, 1983). According to Hayduk, the measurements and methodological strategies used to study interpersonal distance needed refinement, which can be achieved through the use of new technologies. One approach has used head-mounted displays to track the subject's head position and orientation and adapt the view on a virtual environment in the display (Bailenson, Blascovich, Beall, & Loomis, 2003). Another has used virtual characters, which allows for the control of stimuli but reduces the ecological validity, especially when the interactions involve dialogue (Dotsch & Wigboldus, 2008).

These problems can be overcome when full-body motion capture devices are used, because it becomes possible to record the interactions between subjects without hindering other means of expression (e.g., facial movement and speech). In addition, more informative measures can be employed, such as those looking at open or closed postures and body orientation (e.g., Mead, Atrash, & Mataric, 2013).

AMAB can be used to determine interpersonal distances automatically. Let us turn to the handshake example described earlier. In addition to pose differences (Fig. 4), we can also derive measures to describe the distance between two subjects A and B from the motion captured data. One common measure is the distance between the head positions of A and B, observed at the same time (e.g., Bailenson et al., 2003).

Equation 5 is applied with only the head joint H in set J . As the position in world space is required, no position and orientation normalization is applied. In Fig. 5, this head distance is shown for the handshake sequence. Note that Equation 5 also takes into account the height of the subjects' heads. Alternatively, this height can be ignored by calculating the distance $\delta_{A,B}$ between head positions x_{H_A} and x_{H_B} in only the x - and z -plane as $\delta_{A,B} = \sqrt{(x_{H_Ax} - x_{H_Bx})^2 + (x_{H_Az} - x_{H_Bz})^2}$. Instead of looking at the distance between the heads of the subjects, we can also calculate the minimum distance between them by considering all joints. In Fig. 5, this distance is lower, as compared with head distance, especially halfway through the sequence when the subjects extend their hands toward each other and perform the handshake. This measure is useful for conversations, when the interpersonal space is regulated by the hands. A third measure can be obtained by looking at the orientation of both subjects. This could be used to analyze whether subjects are facing each other. The measure can be obtained by applying Equation 2 and comparing the values of θ between subjects. This difference is shown in Fig. 5, with low values corresponding to a similar orientation. There is a decrease in orientation difference after the handshake, due to one subject turning while walking away while the other walks backward. These measures of interpersonal distance can give more insight in the displayed behavior and allow one to go beyond the traditional measures.

Comparisons across body parts

Although much of what we have discussed in relation to AMAB involves treating the body as a whole and calculating differences across full-body poses, it is also possible to isolate joint positions and make pairwise comparisons. For example, researchers interested in gesture may wish to differentiate and compare the movements of both hands (e.g., McNeill, 1992).

One prominent line of research is concerned with the relationship between body pose and perceived affect (Kleinsmith & Bianchi-Berthouze, 2013). In a typical study of this issue, subjects are shown stimuli of manipulated body poses and are asked to assign affective labels. These manipulated poses involve deliberate variation in the positions or orientations of body parts, such as a right elbow bend of 45° . The obtained ratings are then used to associate different patterns in position and orientation of body parts with different ratings of perceived affective state. While these stimuli can be varied systematically, their ecological validity is often lower as the body parts are arranged in an unnatural manner. To this end, researchers have employed motion capture devices to record full-body poses while eliciting emotions and, thereby, capturing body poses that correlate with genuine emotions (e.g., Kleinsmith, Bianchi-Berthouze, & Steed, 2011).

AMAB could be used as a standardized approach to conducting these studies. In addition to the analysis of static body poses, the use of motion capture devices also allows for the analysis of dynamic aspects of affect by measuring the velocity or amount of movement for individual body parts, using Equations 6 and 7. Using the same methodology, one could examine emotion contagion by analyzing to what extent subjects assume a displayed body pose that is associated to a certain emotion. Such studies might reveal different patterns for individual body parts.

Comparisons across people

Often, one is interested in comparing body movements of multiple people—for example, subjects performing the same task at different moments in time, such as performing gestures. Alternatively, one can look at body movement of multiple interacting subjects at the same moment in time, such as occurs in studies on pedestrian avoidance in crowded places and on turn-taking in interactions. The example that will be explained in more detail here is the occurrence of behavioral mimicry in interactions.

Nonconscious mimicry is the automatic tendency to imitate the behaviors of other people, including poses, gestures, mannerisms, speech rates, and facial expressions (Chartrand & Bargh, 1999; Stel, van Dijk, & Olivier, 2009) at the same time or within a short time window (Chartrand & Lakin, 2013). Increased levels of mimicry facilitate smooth interactions and foster liking (Chartrand & Bargh, 1999), and recent research has focused on the moderators and consequences of behavioral mimicry (Chartrand & Lakin, 2013). So far, in most studies, manually coded events from video recordings are used to measure behavioral mimicry (e.g., Stel et al., 2009). Besides issues with the subjective and time-consuming nature of the task (Scherer & Ekman, 1982), these comparisons are usually only made between isolated behaviors (e.g., face touching and gesturing). This excludes quantitative analyses of the form, magnitude, and direction of the behavior. Methods that directly measure synchrony from video recordings (e.g., Paxton & Dale, 2013) allow for such a quantitative analysis but are strongly influenced by nuisance factors such as camera viewpoint, illumination, and type of clothing. The motion capture devices described in this article do not suffer from these drawbacks.

The AMAB methodology can be used to numerically analyze the amount of mimicry between 2 subjects A and B, by looking at their poses or their motion. When using manually coded videos, the occurrence of individual behaviors (e.g., posture shift or head nod) is typically rather low, which requires the use of fairly large time intervals. In contrast, the frame rate of the body motion recordings is typically high, which enables analysis of mimicry at a much finer time scale. To make sure only the pose, not absolute position, is taken into

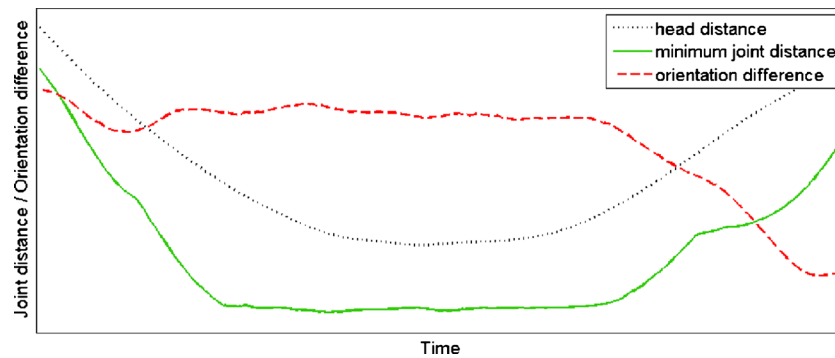


Fig. 5 Different interpersonal distance measures over time

account, the poses of both subjects are normalized using Equation 1. Since interacting subjects typically face each other, poses are also normalized for rotation using Equation 3. This ensures that both subjects have similar positions and facing directions (cf. Fig. 3). Additionally, one might left–right mirror the pose of one of the subjects, to make the direction of the movements of both subjects similar. As both subjects are rotated to face the positive x -axis, the z -values of all joints j in J of one of the subjects can be negated: $x_{jz} = -x_{jz}$. Mimicry can also be operationalized using body motion by analyzing the velocity of the body. The screened data of both subjects can subsequently be compared using windowed cross-correlation (e.g., Paxton & Dale, 2013) or based on spectral methods (e.g., Oullier, de Guzman, Jantzen, Lagarde, & Kelso, 2007) with pairwise pose distances calculated using Equation 5.

Conclusions and future research

We have introduced a set of standards and techniques for studying nonverbal behavior as measured with full-body motion tracking technology. The increasingly wider availability and applicability of these devices provide opportunities for psychologists working on nonverbal behavior, but it is important that the complexity they bring is handled in a sensible and consistent manner. The approach we propose is one possible standardized methodology that addresses the automatic measurement and quantitative analysis of full-body motion for a broad range of applications and research questions. It is worth stressing the word *quantitative* in that description, since AMAB neither gives a qualitative description of the recorded body motion (e.g., the left arm is moving upward) nor provides an interpretation of the movement (e.g., the person is reaching). While the former could be obtained by defining rules on the measured motion (Dael et al., 2012), the latter requires knowledge of the context in which the movement is performed. This depends strongly on the specific research question and experiment setup and is, therefore, outside the scope of AMAB.

While this article has focused on hypothesis-driven research questions, AMAB can also be used in explorative,

data-driven research. Instead of using one or a few dependent variables, a large number of variables (i.e., features) can be derived from the body motion measurements. The automatic recording and subsequent screening of the data provide an excellent starting point for the calculation of these features for use in statistical analyses. In such a pattern recognition approach, a computational model is derived from a subset of the labeled data. Such a model predicts the label given a set of body motion features. For example, Kleinsmith et al. (2011) determined the affective category associated to a body pose, described as a set of joint angles. Krishnan et al. (2009) classified hand movements from a number of features derived from accelerometers attached to the hands. While these works were aimed at automatic recognition, data-driven research can also be used to explore the contribution of measured body motion in more detail.

There are several extensions possible in the development of the AMAB methodology. One potential extension stems from the fact that AMAB does not describe the motion in terms of qualitative labels per limb, or combinations of limbs, as in the BAP coding scheme (Dael et al., 2012). It will be useful to develop a methodology for automatic translation of the quantitative representation of body movement data into qualitative form. Such a translation can be implemented by means of manually crafted rules that take the movement of a single or multiple limbs and assign a label from a set of codes. Alternatively, pattern recognition methods can be used to automatically learn such a mapping from labeled data. However, this is challenging, due to the large variation in possible body movements.

A second set of potential extensions concern the representation and analysis of the body motion of people. For example, it will be useful to develop analyses that go beyond the pairwise comparisons presented in the present article and provide opportunities to objectively study coordination between subjects in terms of temporal and spatial patterns of their body movement. In particular, we also foresee extensions in the analysis of group behavior. It may also be valuable to address interpersonal differences in body movement in a more explicit way than addressed here. Dealing with variations in the amount and type of movements could lead to a notion of

baseline behavior, which is instrumental in many research questions. This further increases the applicability of AMAB.

Acknowledgment The authors acknowledge financial support from the Dutch programme COMMIT and from the EU FP7 network of excellence SSPNet.

References

- Altorfer, A., Jossen, S., Wurmle, O., Kasermann, M.-L., Foppa, K., & Zimmermann, H. (2000). Measurement and meaning of head movements in everyday face-to-face communicative interaction. *Behavior Research Methods, Instruments, & Computers*, *32*, 17–32. doi:10.3758/BF03200785
- Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, *29*, 819–833. doi:10.1177/0146167203029007002
- Bente, G. (1989). Facilities for the graphical computer simulation of head and body movements. *Behavior Research Methods, Instruments, & Computers*, *21*, 455–462.
- Bente, G., Petersen, A., Krämer, N. C., & De Ruiter, J. P. (2001). Transcript-based computer animation of movement: Evaluating a new tool for nonverbal behavior research. *Behavior Research Methods, Instruments, & Computers*, *33*, 303–310. doi:10.3758/BF03195383
- Bente, G., Senokozlieva, M., Pennig, S., Al-Issa, A., & Fischer, O. (2008). Deciphering the secret code: A new methodology for the cross-cultural analysis of nonverbal behavior. *Behavior Research Methods*, *40*, 269–277. doi:10.3758/BRM.40.1.269
- Bezodis, N. E., Salo, A. I. T., & Trewartha, G. (2010). Choice of sprint start performance measure affects the performance-based ranking within a group of sprinters: Which is the most appropriate measure? *Sports Biomechanics*, *9*, 258–269. doi:10.1080/14763141.2010.538713
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910. doi:10.1037/0022-3514.76.6.893
- Chartrand, T. L., & Lakin, J. L. (2013). The antecedents and consequences of human behavioral mimicry. *Annual Review of Psychology*, *64*, 285–308. doi:10.1146/annurev-psych-113011-143754
- Dael, N., Mortillaro, M., & Scherer, K. R. (2012). The body action and posture coding system (BAP): Development and reliability. *Journal of Nonverbal Behavior*, *36*, 97–121. doi:10.1007/s10919-012-0130-0
- Dakin, C. J., Luu, B. L., van den Doel, K., Inglis, J. T., & Blouin, J. (2010). Frequency-specific modulation of vestibular-evoked sway responses in humans. *Journal of Neurophysiology*, *103*, 1048–1056. doi:10.1152/jn.00881.2009
- Doron, C., Beattie, G., & Shovelton, S. (2010). Nonverbal indicators of deception: How iconic gestures reveal thoughts that cannot be suppressed. *Semiotica*, 133–174. doi:10.1515/semi.2010.055
- Dotsch, R., & Wigboldus, D. H. J. (2008). Virtual prejudice. *Journal of Experimental Social Psychology*, *44*, 1194–1198. doi:10.1016/j.jesp.2008.03.003
- Ekman, P. (1965). Communication through nonverbal behavior: A source of information about an interpersonal relationship. In S. S. Tomkins & C. Izard (Eds.), *Affect, cognition, and personality* (pp. 390–442). New York: Springer.
- Feese, S., Amrich, B., Tröster, G., Meyer, B., & Jonas, K. (2012). Quantifying behavioral mimicry by automatic detection of nonverbal cues from body motion. In A. Nijholt, A. Vinciarelli, B. Schüller, & M. Smith (Eds.), *International conference on social computing* (pp. 520–525). Piscataway, NJ: IEEE. doi:10.1109/SocialCom-PASSAT.2012.48
- Frey, S., & Von Cranach, M. (1973). A method for the assessment of body movement variability. In M. von Cranach & I. Vine (Eds.), *Social communication and movement* (pp. 389–418). New York, NY: Academic.
- Grammer, K., Kruck, K. B., & Magnusson, M. S. (1998). The courtship dance: Patterns of nonverbal synchronization in opposite-sex encounters. *Journal of Nonverbal Behavior*, *22*(1), 3–29. doi:10.1023/A:1022986608835
- Hall, E. T. (1966). *The hidden dimension*. New York, NY: Doubleday.
- Hayduk, L. A. (1983). Personal space: Where we now stand. *Psychological Bulletin*, *94*, 293–335. doi:10.1037/0033-2909.94.2.293
- Hirsbrunner, H. P., Frey, S., & Crawford, R. (1987). Movement in human interaction: Description, parameter formation, and analysis. In A. Siegman & S. Feldstein (Eds.), *Nonverbal behavior and communication* (pp. 99–140). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kleinsmith, A., & Bianchi-Berthouze, N. (2013). Affective body expression perception and recognition: A survey. *IEEE Transactions on Affective Computing*, *4*, 15–33. doi:10.1109/T-AFFC.2012.16
- Kleinsmith, A., Bianchi-Berthouze, N., & Steed, A. (2011). Automatic recognition of non-acted affective postures. *IEEE Transactions on Systems, Man, and Cybernetics - Part B*, *41*, 1027–1038. doi:10.1109/TSMCB.2010.2103557
- Krishnan, N. C., Juillard, C., Colby, D., & Panchanathan, S. (2009). Recognition of hand movements using wearable accelerometers. *Journal of Ambient Intelligence and Smart Environments*, *1*, 143–155. doi:10.3233/AIS-2009-0019
- Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, *41*, 841–849. doi:10.3758/BRM.41.3.841
- McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, *92*, 350–371. doi:10.1037/0033-295X.92.3.350
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- Mead, R., Atrash, A., & Mataric, M. J. (2013). Automated proxemic feature extraction and behavior recognition: Applications in human-robot interaction. *International Journal of Social Robotics*, *5*, 367–378. doi:10.1007/s12369-013-0189-8
- Oullier, O., de Guzman, G. C., Jantzen, K. J., Lagarde, J., & Kelso, J. A. S. (2007). Social coordination dynamics: Measuring human bonding. *Social Neuroscience*, *3*, 178–192. doi:10.1080/17470910701563392
- Paxton, A., & Dale, R. (2013). Frame-differencing methods for measuring bodily synchrony in conversation. *Behavior Research Methods*, *45*, 329–343. doi:10.3758/s13428-012-0249-2
- Poppe, R. (2007). Vision-based human motion analysis: An overview. *Computer Vision and Image Understanding*, *108*, 4–18. doi:10.1016/j.cviu.2006.10.016
- Scherer, K. R., & Ekman, P. (1982). Methodological issues in studying nonverbal behavior. In K. R. Scherer & P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research* (pp. 45–135). New York: Cambridge University Press.
- Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A. W., Finocchio, M., Blake, A., ... Moore, R. (2013). Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, *56*, 116–124. doi:10.1145/2398356.2398381
- Slawinski, J., Dumas, R., Cheze, L., Ontanon, G., Miller, C., & Mazure-Bonnefoy, A. (2013). Effect of postural changes on 3D joint angular velocity during starting block phase. *Journal of Sports Sciences*, *31*, 256–263. doi:10.1080/02640414.2012.729076
- Stel, M., van Dijk, E., & Olivier, E. (2009). You want to know the truth? Then don't mimic! *Psychological Science*, *20*, 693–699. doi:10.1111/j.1467-9280.2009.02350.x
- Vick, S. J., Waller, B. M., Parr, L. A., Smith Pasqualini, M. C., & Bard, K. A. (2006). A cross-species comparison of facial morphology and movement in humans and chimpanzees using the Facial Action Coding System (FACS)". *Journal of Nonverbal Behavior*, *31*, 1–20. doi:10.1007/s10919-006-0017-z