Article

Is Markerless More or Less? Comparing a Smartphone Computer Vision Method for Equine Lameness Assessment to Multi-Camera Motion Capture

Felix Järemo Lawin [1], Anna Byström [2], Christoffer Roepstorff [1], Marie Rhodin [2], Mattias Almlöf [1], Mudith Silva [1], Pia Haubro Andersen [2], Hedvig Kjellström [3] and Elin Hernlund [1,2,*]

1 Sleip AI, Birger Jarlsgatan 58, 11426 Stockholm, Sweden
2 Department of Anatomy, Physiology and Biochemistry, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden
3 KTH Royal Institute of Technology, Division of Robotics, Perception and Learning, 10044 Stockholm, Sweden
* Correspondence: elin.hernlund@slu.se; Tel.: +46-18-672142

Simple Summary: Lameness, an alteration of the gait due to pain or dysfunction of the locomotor system, is the most common disease symptom in horses. Yet, it is difficult for veterinarians to correctly assess by visual inspection. Objective tools that can aid clinical decision making and provide early disease detection through sensitive lameness measurements are needed. In this study, we describe how an AI-powered measurement tool on a smartphone can detect lameness in horses without the need to mount equipment on the horse. We compare it to a state-of-the-art multi-camera motion capture system by simultaneous, synchronised recordings from both systems. The mean difference between the systems' output of lameness metrics was below 2.2 mm. Therefore, we conclude that the smartphone measurement tool can detect lameness at relevant levels with easy-of-use for the veterinarian.

Abstract: Computer vision is a subcategory of artificial intelligence focused on extraction of information from images and video. It provides a compelling new means for objective orthopaedic gait assessment in horses using accessible hardware, such as a smartphone, for markerless motion analysis. This study aimed to explore the lameness assessment capacity of a smartphone single camera (SC) markerless computer vision application by comparing measurements of the vertical motion of the head and pelvis to an optical motion capture multi-camera (MC) system using skin attached reflective markers. Twenty-five horses were recorded with a smartphone (60 Hz) and a 13 camera MC-system (200 Hz) while trotting two times back and forth on a 30 m runway. The smartphone video was processed using artificial neural networks detecting the horse's direction, action and motion of body segments. After filtering, the vertical displacement curves from the head and pelvis were synchronised between systems using cross-correlation. This rendered 655 and 404 matching stride segmented curves for the head and pelvis respectively. From the stride segmented vertical displacement signals, differences between the two minima (MinDiff) and the two maxima (MaxDiff) respectively per stride were compared between the systems. Trial mean difference between systems was 2.2 mm (range 0.0–8.7 mm) for head and 2.2 mm (range 0.0–6.5 mm) for pelvis. Within-trial standard deviations ranged between 3.1–28.1 mm for MC and between 3.6–26.2 mm for SC. The ease of use and good agreement with MC indicate that the SC application is a promising tool for detecting clinically relevant levels of asymmetry in horses, enabling frequent and convenient gait monitoring over time.

Keywords: monocular motion analysis; objective lameness assessment; equine orthopaedics; animal pose estimation; optical motion capture

1. Introduction

Objective measurement of a horse's motion at the trot has become an important part of the diagnostic procedures performed during clinical lameness investigation. These measurements, which have been used in clinical practice for more than a decade, trace the vertical displacement of axial body segments: the head, the pelvis and sometimes the withers. Using reflective markers or inertial sensors attached to a point on each body segment, a time series signal is generated. In trot, the vertical displacement signal takes the shape of a sinusoidal double wave from each stride and it is the position of the two peaks and valleys of this signal which are used for lameness analysis.

The degree of asymmetry in the vertical displacement signal i.e., the difference between the two peaks and valleys respectively is known to indicate asymmetric loading of the left versus right limb during the midstance and the push-off phases of the stride [1–3]. Measurements of these asymmetries provide the veterinarian with high-resolution data that help overcome the limited time resolution of the human visual system [4,5]. These objective data seem crucial for quality control of the clinical procedure, since subjective lameness assessment has been shown to have moderate to low agreement between veterinarians [6,7] and is affected by expectation bias [8]. The metrics derived from objective motion analysis show high sensitivity for single-limb lameness, acting as early indicators to detect asymmetric loading of the limbs [9]. However, the specificity for lameness on a population level is less clear. Motion asymmetries are commonly observed in cross-sectional studies of different horse populations, such as Warmblood riding horses [10,11], Thoroughbred race horses [12], working Polo horses [13], elite eventing horses [14], endurance horses [15] and young Standardbred trotters [16], with a prevalence ranging between 50–90 percent. Although these asymmetries are of the same magnitude as in horses investigated for lameness in a clinical setting [17], it is currently unknown if these asymmetry levels indicate that a large proportion of horses in training are lame or if the asymmetries can be explained by other factors, such as laterality. A key approach to the further investigation of this issue is to perform longitudinal monitoring of individual horses over time. For this to be possible, a reliable, ease of use and low-cost measuring system is required.

Several motion analysis systems have been developed for clinical use based on inertial measurement units (IMUs) [18–20]. Also available is a multi-camera marker-based motion capture (MC) system [21]. This MC-system, is considered to be the gold standard for measuring body segment movement for kinematic gait analysis [22]. It relies on reflective markers attached to a horse's body. These markers are detected and tracked by a set of cameras that are geometrically calibrated and temporally synchronized. By using the synchronized tracks of the marker positions in the camera images, MC-systems reconstruct the 3D coordinates using multi-view triangulation. It has been shown that under favourable conditions, the accuracy of the computed estimates of the 3D positions over time is less than a millimetre for the MC-systems [23]. However, placing markers on the horse are resource and time-consuming in a clinical situation, and the equipment is a substantial financial investment for a veterinary practice. This impedes the system's large-scale clinical and scientific use.

During the previous decade, the field of computer vision was revolutionized by methods based on deep neural networks [24–26]. These networks are computer algorithms that consist of multi-layered (referred to as deep) compositions of parametric functions that can be trained on large datasets to perform classification and regression tasks. Deep learning has demonstrated increased robustness to differing scenarios, light conditions, and noise levels compared to traditional computer vision methods. Estimation of poses of the human body from images has been enabled by deep learning and as a result of this development, it has become possible to perform motion analysis from video, e.g., from a smartphone camera [27]. Recent works on horse lameness classification [28,29] have demonstrated a progressive movement towards the application of computer vision and deep learning within objective motion analysis. However, a binary disease classifier of "lame" versus "not lame" is a difficult approach for clinical use, given that lameness is often not a binary state

and that the deep learning algorithms act in a black box manner, rendering distrust from medical professionals [30]. Instead, providing a clinician with computer vision derived metrics to support medical decision making is a more implementable approach. But until now, the methodological accuracy of deep neural networks for quantification of clinically used lameness metrics has not been investigated.

In this work, we validate a new single-camera markerless (SC) system designed for equine lameness assessment which uses images from a smartphone camera video stream. To achieve robust detection and tracking, the system employs a series of neural networks. These networks were trained to detect and track the pelvis, head and hooves in video streams of trotting horses. The system also detects the trotting direction of the horse, away from or towards the camera, to determine which parts of the horse are visible for measuring. The network designs were inspired by previously proposed methods for object detection [28,31] and segmentation [32,33]. Unlike the MC-system, the deep neural networks of the SC-system do not require that markers are placed on the horse. Instead, the networks learn to detect the points of interest on the horse's body visible in the images through training on large datasets.

The specific aim of the study was to compare this new markerless smartphone system to a state-of-the-art multi-camera marker-based system with respect to waveform similarity of the derived vertical displacement signals and the limits of agreement for their extracted lameness metrics.

2. Materials and Methods

2.1. Study Protocol

Twenty-five horses were recorded as they underwent motion analysis at the orthopedic gait laboratory situated in the Equine Clinic of the University Animal Hospital in Uppsala, Sweden. The recordings were performed simultaneously with a multi-camera marker-based motion capture system and a single smartphone camera markerless system. The experimental setup is illustrated in Figure 1. The study subjects were a convenience sample selected from the horses visiting the clinic during the 10-day data collection period, without any exclusion criteria. The horses were of different breeds, sizes (range 128 to 180 cm to the withers) and colours (black, bay, chestnut, grey). All owners gave their written informed consent to participate. The study did not in any way alter the clinical procedures or add physical manipulation of the animals. Hence, no ethical approval was required according to the national animal ethics legislation.

2.2. Data Collection

During data collection, each horse was guided by a handler (the owner, or a researcher running at the horse's left side) to trot at least two times back and forth on a 30 m concrete runway in a corridor. The horses were jogged at the handler's preferred speed.

We employed a MC-system with 13 cameras (Qualisys AB, Motion Capture Systems). The MC cameras were placed $\approx$ 4 m over the ground and in a manner such that the union of the field of view of each camera covered as much of the runway as possible while still maintaining sufficient overlap between neighboring cameras to perform tracking and triangulation of the marker positions. The recording rate was set to 200 Hz. We attached spherical reflective markers in the median plane over the poll on the horse's head and between the tubera sacrale of the pelvis, allowing the MC-system to detect and track the markers over time. Additional markers were placed, but not used in this study (see Figure 2).

The smartphone (iPhone12 Pro Max) was placed $\approx$ 1.6 m above the ground on a tripod facing the direction of the horse trot recording 4k video (2160 $\times$ 3840 pixels) at 60 Hz. The video streams recorded were input to the SC-system. Example frames from the smartphone camera recordings are shown in Figure 2.
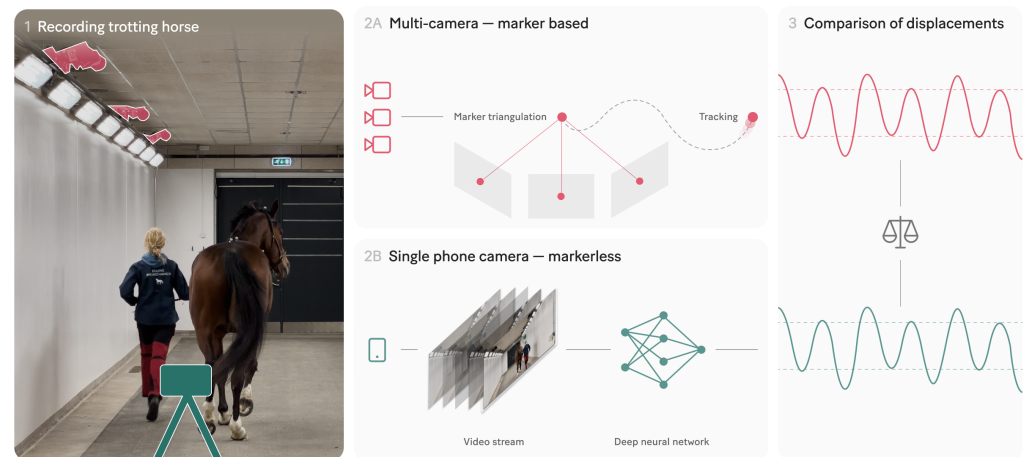
Figure 1. Illustration of the experimental setup. We recorded horses trotting back and forth in a corridor with a multi-camera system and a single smartphone camera. The multi-camera system detected and tracked reflective markers attached to the horse. Marker positions were triangulated into 3D coordinates from which vertical displacement curves were extracted. The single-camera system used deep neural networks to predict the vertical displacement curves of the head and pelvis from the images in the smartphone video stream. We then compared the displacement curves from the two systems using normalised peak and valley differences.

For each recording, the MC-system and the smartphone camera were triggered at approximately the same time such that the data streams from both systems would cover the same sequence of events.

Figure 2. Example images from data set (horse 21) taken from the video recorded by the markerless single-camera system (SC). Attached to the horse's skin by double adhesive tape are the spherical reflective markers used by the multi-camera marker-based system (MC) for tracking head and pelvic motion. The markers on the poll and between the tubera sacrale were used.

2.3. Signal Extraction

The MC-system software (Qualisys Track Manager—QTM, Qualisys AB) automatically tracked the reflective markers and generated 3D coordinates corresponding to the positions of the markers in each frame. The 3D marker coordinates from QTM were exported to .mat files (MATLAB). From these coordinates, vertical displacements were extracted for each frame, one for the head marker and one for the pelvis marker. This resulted in two vertical

displacement signals (VDS), $y_{mc}^{head}(t)$ and $y_{mc}^{pelvis}(t)$ for head and pelvis respectively, where t is time in seconds.

For the SC-system, deep neural networks were applied (software of Sleip AI AB) on the input video stream from the smartphone camera. The deep neural networks were trained to output the pixel coordinates of horse body parts for each frame of the video. The training material for the deep neural networks contained horses of many different coat colours and varying conformation, but none had physical markers attached to the skin. Head ($y_{sc}^{head}(t)$) and pelvis ($y_{sc}^{pelvis}(t)$) VDS was calculated from the pixel coordinates. Additionally, the VDS of all four hooves were extracted from the SC-system for stride splitting purposes (see Section 2.4). Both the MC-system and the SC-system data were further processed using custom written python scripts.

Note that the pelvis was visible to the SC-system only when the horse was trotting away from the camera, while the head was mostly visible in both directions. Consequently, SC generally produced data from a higher number of strides with head tracking than strides with pelvic tracking. The following analyses only included strides with data matched from both systems for the body segment in question (head and pelvis). Horses were removed from the dataset where less than 10 matching strides were available for either head or pelvis since we deemed that insufficient for statistical relevance.

2.4. Stride Split and Signal Filtering

The recorded data contained noise, due to measurement errors and because the horses seldom trot in a consistent manner throughout a trot-up. This noise was present for both the MC-system and the SC-system data. Thus, the VDS had to be band-pass filtered in order to remove the noise without affecting the frequency content of the signal that related to movement asymmetries and lameness [34].

In order to perform the VDS filtering described in [34], the within horse mean stride frequency of a measurement was needed. This was estimated by extracting strides from the hoof VDS of the SC-system. Firstly, we performed a pre band-pass filtering of the hoof VDS to remove trends and high-frequency noise. Specifically, a 7th order Butterworth digital filter with a lower bound cut-off frequency of 0.6 Hz and an upper bound cut-off frequency of 2.2 Hz. This allowed us to determine in which time intervals the left hoof was above the right hoof and vice versa, ultimately enabling the classification of left and right strides.

Next, the lengths of the intervals were used to compute the stride frequency, which in turn was used to set the bounds of the 10th order Butterworth band-pass filter applied to the VDS of the pelvis and head. Specifically, we set the lower bound to 0.75 times the stride frequency, to not alter frequency content related to the movement asymmetry [34]. Similarly, the upper bound was set to 2.42 times the stride frequency to not attenuate the frequency content related to the symmetric movement, while omitting higher frequency content and noise.

2.5. Signal Synchronization

Since the MC and SC recordings were triggered manually, the extracted signals were not adequately synchronized in time. To synchronize the vertical displacement signals we computed cross-correlations to find the relative time shift t_{shift} that solved the following maximization problem:

$$t_{shift} = \arg\max_t \left((y_{mc}^{pelvis} \star y_{sc}^{pelvis})(t) + (y_{mc}^{head} \star y_{sc}^{head})(t) \right). \tag{1}$$

Here $\star$ denotes the correlation operator. The signals were band-pass filtered according to Section 2.4 before synchronization.

2.6. Asymmetry Quantification

In this section, we introduce a number of definitions that we use in the remaining parts of the paper. First, we define a stride segment as a section of the VDS corresponding to a

time interval of a full stride. To extract the stride segment we utilize the extreme values (see the illustration in Figure 3). Tracing of the vertical displacement of the horse's head or pelvis while it trots yields a sine-shaped signal as depicted in Figure 3. The valleys (local minima) and peaks (local maxima) of the signal are associated with the vertical forces generated during impact (the more force, the lower the valley position) and the relative timing of horizontal and vertical forces during the propulsive phase of the stride respectively (less push-off rendering a lower peak position). Differences between consecutive peaks or valleys can be quantified into an asymmetry index [22], which can be used as an indicator of lameness severity. Depending on the measurement technique and due to variation in horse size, these values may need to be normalised to be comparable [35,36]. Therefore, normalisation of these values to the range of motion (R) is used in the SC-system to obtain values that are more independent of horse size.
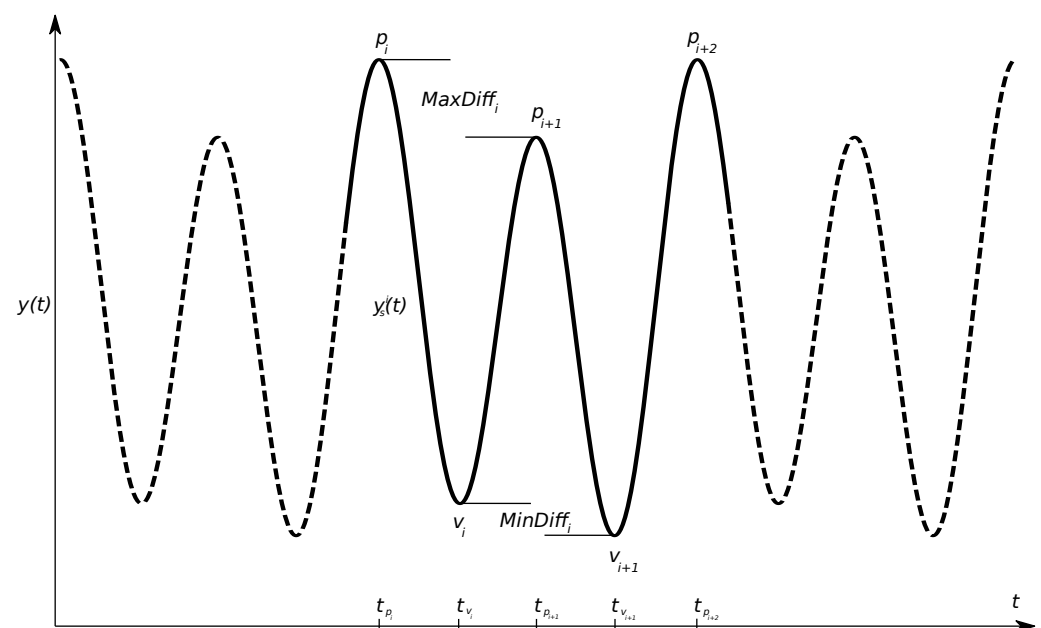
Figure 3. Example of the vertical displacement signal from head or pelvis, with local minima denoted with v for valley and local maxima with p for peak. Metrics for lameness quantification are calculated as the difference between the two minima (MinDiff) and the two maxima (MaxDiff) per stride. A stride segment $y_s^i(t)$ was defined as the section of $y(t)$, starting at t_{p_i} and ending at $t_{p_{i+2}}$.

2.6.1. Extraction of Valleys and Peaks

To find the extreme values within each stride segment (the VDS from one stride), consecutive data points were compared to find points at which the derivative of the VDS was zero. Let $y(t)$ be the VDS value at time t. We defined the peaks $p_i = y(t_{p_i}), i = 1, 2, \ldots$, as the local maximum values and the valleys $v_j = y(t_{v_j}), j = 1, 2, \ldots$, as the local minimum values. Further, we assumed that $t_{p_i} < t_{p_{i+1}}$ and $t_{v_j} < t_{v_{j+1}}$. We extracted a sequence of consecutive peaks and valleys $p_i, p_{i+1}, v_j, v_{j+1}$ such that $t_{p_i} < t_{v_j} < t_{p_{i+1}} < t_{v_{j+1}} < t_{p_{i+2}}$.

In horses showing moderate to severe lameness, the changes in motion asymmetry can cause extreme asymmetries in $y(t)$. In these cases, local extreme values might be canceled out, interrupting the assumed stride pattern of two peaks and two valleys in sequence. Instead, a single peak or valley signal pattern occurs. To handle this, we implemented a robust extreme value extraction method. The reasoning behind this method is that $y(t)$ contains two dominant harmonics [1]. The first harmonic corresponds to the stride frequency, thus contributing to the asymmetry of the signal. The second harmonic corresponds to twice the stride frequency and should dominate $y(t)$ if the horse is healthy. In our approach, we extracted extreme values based on the curve shape of the second

harmonic. We first removed the asymmetric component of $y(t)$ by performing high-pass filtering with a frequency bound higher than the first harmonic. Next, we employed local extreme value extraction on the high-pass filtered signal. Finally, we refined the selection by selecting the extreme values from the original $y(t)$ within 50 ms of the estimated value. As a result, we were able to estimate normalised peak/valley differences at any degree of lameness.

2.6.2. Normalised Differences for Valleys and Peaks

From the extracted stride peaks we computed local extreme value differences per stride i, consisting of the following scalars,

$$\text{MinDiff}_i = v_{i+1} - v_i \tag{2}$$
$$\text{MaxDiff}_i = p_i - p_{i+1}. \tag{3}$$

These values provide information about the asymmetries between the right and left leg impact and push-off. However, since different horses have different amplitudes in their vertical displacement trajectories, these values are not comparable. In this work, we instead use the normalised extreme value differences (NEVd), which show more independence from the scale of the vertical displacement. The NEVd-values were computed using the following operations,

$$V_i = \frac{\text{MinDiff}_i}{R_i} \tag{4}$$
$$P_i = \frac{\text{MaxDiff}_i}{R_i} \tag{5}$$
$$\text{where } R_i = \max(p_i, p_{i+1}) - \min(v_i, v_{i+1}).$$

Here, normalization was performed by division by the range-of-motion R_i. Thus, the NEVd-values measures asymmetry as a rate of R_i.

2.6.3. Outlier Removal

To remove occasional strides with erroneous measurements from the analysis, we performed a series of outlier removal steps.

While the noise in the vertical displacement signals is partly suppressed by band-pass filtering, the signal quality becomes inadequate for asymmetry analysis when the noise is too prevalent. Therefore, we first removed strides where the stride segments contained a substantial amount of high-frequency noise. A stride segment was deemed to contain too much high frequency noise when the majority of the frequency content amplitudes were found above 10 Hz in at least a quarter of a stride interval.

As a second step, we performed a linear discriminant analysis (LDA). In addition to MinDiff_i and MaxDiff_i from Equation (4), we used the peak valley differences $p_i - v_i$ and $p_{i+1} - v_{i+1}$ to represent each stride as features. We then removed NEVd-values from the analysis that corresponded to strides that were considered outliers in the LDA.

2.7. System Comparisons

We compared the MC and SC systems on the data set of trotting horses described in Section 2.2. For each recorded sample, we used MC and SC to generate vertical displacement signals. From these, we extracted and compared stride segments and NEVd-values between the two systems. The following sections detail the implementation and setup of the comparison.

2.7.1. Comparison Metrics

To compare the extracted asymmetry indices between the MC and SC systems the following deviations were calculated from the synchronized NEVd-vales,

$$\Delta V^i = \text{MinDiff}_i^{sc} - \text{MinDiff}_i^{mc} \tag{6}$$

$$\Delta P^i = \text{MaxDiff}_i^{sc} - \text{MaxDiff}_i^{mc} \, . \tag{7}$$

To recover an estimate of geometric deviation, we multiplied ΔV^i and ΔP^i with the range of motion $R_{i,mc}$ computed from the NEVd-values of the MC signal.

In practice, lameness indication is deduced from the trial mean of the NEVd-values $\bar{V} = 1/N \sum_i^N \text{MinDiff}_i$ and $\bar{P} = 1/N \sum_i^N \text{MaxDiff}_i$, where N is the number of strides in the trial after the outlier removal in Section 2.6.3. We computed the trial mean deviations as,

$$\Delta \bar{V} = \bar{V}^{sc} - \bar{V}^{mc} \tag{8}$$

$$\Delta \bar{P} = \bar{P}^{sc} - \bar{P}^{mc} \, . \tag{9}$$

We further scaled $\Delta \bar{V}$ and $\Delta \bar{P}$ with the mean range of motion $\bar{R}_{mc} = 1/N \sum_i^N R_{i,mc}$ to estimate the geometric deviation.

2.7.2. Statistical Analysis

Bland-Altman analysis [37] was used to evaluate the statistical agreement between the MC and SC-systems for head and pelvis NEVd-values. The Bland-Altman analysis was subdivided into deviations for MinDiff and MaxDiff and was performed both on trial and stride level.

In addition to the comparison of the NEVd-values, we compared the shapes of the band pass filtered vertical displacement signals using the root mean square deviations (RMSD),

$$RMSD = \sqrt{\frac{\sum_{m=1}^{M_i} \left(R_{i,sc}^{mc} \cdot y_{sc}(t_{m,sc}^i) - y_{mc}(t_{m,mc}^i) \right)^2}{M_i}} \tag{10}$$

$$\text{where } R_{i,sc}^{mc} = \frac{R_{i,mc}}{R_{i,sc}} \, .$$

Here, $M = 23$ is the number of trials, i.e number of horses in the dataset. $t_{m,sc}^i$ and $t_{m,mc}^i$ are equally spaced points in time for sampling the vertical displacement signals corresponding to the ith synchronized stride, i.e $t_{1,sc}^i = p_{i,sc}$, $t_{M,sc}^i = p_{i+2,sc}$, $t_{1,mc}^i = p_{i,mc}$ and $t_{M_i,mc}^i = p_{i+2,mc}$. Note that we scaled the y_{sc} samples to the geometric scale of y_{mc}. Further, since y_{sc} and y_{mc} have different frame rates, we aligned and re-sampled the signals with linear interpolation before applying (10).

3. Results

In this section, we outline the results from the experiments described in Section 2.7. In total, the results below were generated from 23 of the 25 horses in the initial data set, after the exclusion of two horses with less than 10 synchronized strides. From the included horses we extracted a total of 655 stride observations of the head motion and 404 stride observations of the pelvic motion. Descriptive statistics of the head measurements from the 23 trials can be found in Table 1 and the pelvis measurement in Table 2.

Table 1. Descriptive statistics for head measurements from the 23 included horses showing number of matched strides per trial (N) and the mean trial deviations between the two systems for the valley values $\Delta\bar{V}$ (MinDiff) and peak values $\Delta\bar{P}$ (MaxDiff). Also, the actual trial means for the $\bar{V}$ (MinDiff) and the $\bar{P}$ are presented per trial for the multi-camera marker-based (*mc*) and the single-camera markerless (*sc*) systems, followed by their within trial standard deviation (σV and σP). From the *mc*, the trial mean range of motion of the vertical displacement signal is presented.

Horse	N	$\Delta\bar{V}$	$\Delta\bar{P}$	$\bar{V}^{sc}$	$\bar{V}^{mc}$	$\bar{P}^{sc}$	$\bar{P}^{mc}$	σV^{sc}	σV^{mc}	σP^{sc}	σP^{mc}	$\bar{R}_{mc}$
1	16	−0.4	−4.8	−1.5	−1.1	13.1	17.9	15.2	15.4	18.1	16.5	66.6
2	23	−2.1	8.7	−9.4	−7.3	43.0	34.3	12.3	10.4	19.7	20.8	82.2
3	29	2.3	−1.9	−1.6	−3.9	9.5	11.4	17.7	15.0	13.6	14.1	70.4
4	38	4.1	3.4	−39.8	−44.0	−0.8	−4.2	14.3	15.7	14.8	16.8	71.2
5	28	−0.3	2.2	70.0	70.3	−17.5	−19.8	17.4	15.2	17.2	14.5	109.6
6	26	4.4	5.7	−39.2	−43.6	−3.9	−9.6	16.8	19.1	15.8	15.8	68.5
7	19	3.3	−3.1	5.4	2.1	3.3	6.3	13.5	14.2	19.9	21.1	77.5
8	22	−0.6	−3.7	7.1	7.7	8.8	12.5	15.7	11.9	11.0	10.4	73.6
9	29	1.0	−5.0	−39.7	−40.7	15.7	20.6	9.1	10.6	9.5	10.0	70.8
10	36	−0.6	0.6	1.8	2.4	−19.1	−19.8	22.3	22.0	20.7	21.4	95.2
11	27	0.0	0.4	−57.8	−57.9	13.0	12.5	11.5	13.1	16.4	15.2	90.1
12	28	−2.5	1.9	−11.7	−9.2	−18.5	−20.4	14.3	13.5	17.5	14.9	71.0
13	22	−0.0	1.6	62.8	62.8	−3.7	−5.3	8.9	8.2	10.8	10.3	79.8
14	22	0.3	−0.9	−1.6	−1.9	8.7	9.6	9.7	13.8	6.1	10.6	39.0
15	29	1.5	−1.8	27.0	25.5	−14.1	−12.4	13.8	13.1	11.2	11.6	75.2
16	19	−6.8	0.2	22.8	29.6	10.9	10.7	19.1	26.2	18.6	28.1	75.5
17	34	−0.2	−3.7	3.7	3.9	−4.2	−0.5	18.9	17.5	19.6	18.7	95.0
18	24	1.8	−3.1	14.9	13.1	−13.1	−10.1	8.3	7.9	13.6	11.4	57.9
19	35	−0.4	−2.1	0.2	0.6	24.2	26.3	6.9	6.4	8.6	6.0	51.0
20	41	−0.5	1.7	−21.6	−21.0	7.0	5.3	11.9	12.6	9.1	8.5	71.4
21	16	−0.4	1.2	−23.5	−23.0	5.8	4.6	8.4	7.7	6.7	7.4	43.0
22	36	2.2	−1.8	−13.0	−15.2	−1.5	0.3	8.5	8.2	12.3	10.5	50.9
23	56	−4.2	−0.1	−18.2	−14.0	−16.3	−16.2	18.7	17.6	18.8	18.9	74.7
mean	28.5	1.7	2.6	21.5	21.8	12.0	12.6	13.6	13.7	14.3	14.5	72.2

Table 2. Descriptive statistics for pelvic measurements from the 23 included horses showing the number of matched strides per trial (N) and the mean trial deviations between the two systems for the valley values $\Delta\bar{V}$ (MinDiff) and peak values $\Delta\bar{P}$ (MaxDiff). Also, the actual trial means for the $\bar{V}$ (MinDiff) and the $\bar{P}$ are presented per trial for the multi camera marker based (*mc*) and the single camera markerless system (*sc*), followed by their within trial standard deviation (σV and σP). From the *mc*, the trial mean range of motion of the vertical displacement signal is presented.

Horse	N	$\Delta\bar{V}$	$\Delta\bar{P}$	$\bar{V}^{sc}$	$\bar{V}^{mc}$	$\bar{P}^{sc}$	$\bar{P}^{mc}$	σV^{sc}	σV^{mc}	σP^{sc}	σP^{mc}	$\bar{R}_{mc}$
1	13	1.0	−0.1	−0.1	−1.1	6.6	6.6	8.4	6.1	8.2	8.3	83.4
2	15	1.7	4.3	−0.4	−2.1	0.6	−3.8	11.0	8.6	10.0	8.6	82.3
3	16	−1.8	2.1	−0.6	1.2	−2.7	−4.8	6.5	4.0	8.9	5.6	82.6
4	24	4.3	6.5	−15.8	−20.1	21.0	14.5	9.3	7.9	9.6	8.1	92.5
5	17	5.5	−0.5	5.7	0.2	4.3	4.8	5.1	4.3	10.3	11.3	79.3
6	15	4.8	5.9	−12.8	−17.6	27.2	21.3	11.0	7.5	8.9	6.3	92.8
7	17	1.5	0.0	−0.1	−1.5	−5.2	−5.2	6.8	6.9	10.5	8.5	74.8
8	13	1.1	3.1	2.1	1.0	−8.6	−11.6	8.0	8.1	9.2	4.7	76.7
9	15	−0.8	2.2	−13.0	−12.2	−4.3	−6.5	7.1	4.6	6.0	6.0	67.5
10	22	4.6	2.1	9.0	4.4	2.2	0.1	6.5	7.7	15.8	16.0	77.7
11	15	−1.7	0.7	−7.0	−5.3	−11.9	−12.6	6.4	8.3	10.2	9.6	88.0
12	14	−0.9	1.1	−4.5	−3.6	2.6	1.5	6.4	4.8	7.3	7.9	64.7

Table 2. *Cont.*

Horse	N	$\Delta\bar{V}$	$\Delta\bar{P}$	$\bar{V}^{sc}$	$\bar{V}^{mc}$	$\bar{P}^{sc}$	$\bar{P}^{mc}$	σV^{sc}	σV^{mc}	σP^{sc}	σP^{mc}	$\bar{R}_{mc}$
13	11	3.9	2.4	−12.6	−16.5	9.0	6.6	9.7	6.9	8.1	8.5	70.8
14	14	1.3	0.9	−5.1	−6.4	13.0	12.0	5.8	6.7	9.0	8.6	74.8
15	14	−2.3	−3.0	5.8	8.1	−16.9	−13.9	5.7	3.1	6.0	3.3	75.3
16	13	4.2	−1.1	11.8	7.7	−25.8	−24.7	10.5	13.0	13.8	8.7	75.7
17	18	−1.4	−0.8	−10.6	−9.3	−1.4	−0.6	12.2	11.2	11.4	11.5	103.7
18	15	−3.2	−0.3	−2.8	0.4	−13.1	−12.8	5.6	4.6	7.1	3.3	85.5
19	26	1.2	−1.2	−9.6	−10.8	4.4	5.6	3.6	3.4	4.3	4.6	39.6
20	21	−4.9	1.3	−7.2	−2.3	0.1	−1.2	6.5	6.6	11.6	9.0	87.8
21	21	0.8	2.6	−38.0	−38.9	48.5	46.0	7.1	6.8	12.7	8.6	97.7
22	25	−1.5	2.8	−10.2	−8.7	−1.4	−4.2	6.6	5.1	5.5	6.1	67.0
23	30	1.0	0.4	−21.5	−22.6	12.5	12.1	8.9	9.8	8.8	8.0	72.6
mean	17.6	2.4	2.0	9.0	8.8	10.6	10.1	7.6	6.8	9.3	7.9	78.8

We split the comparisons into per-stride and per-trial comparisons. In the per-stride comparisons, we treated each stride as a sample and performed statistical analysis on the stride-based deviations ΔV_i and ΔP_i. In the per-trial comparison, we treated each horse as a sample and performed statistical analysis on deviations computed from the mean NEVd values $\Delta\bar{V}_i$ and $\Delta\bar{P}_i$.

3.1. Per-Stride Comparisons

An example of the time-domain curves from the MC and SC for all strides in a recorded trial (horse 11) are displayed in Figure 4. The displayed stride segments of the two systems show high resemblance, resulting in similar conclusions on lameness diagnosis. We provide more examples from the experiment in Appendix A.

The Bland-Altman analysis for head and pelvis lameness metrics is illustrated in Figures 5 and 6 respectively. The deviations are similar for both valley and peak differences and generally higher for head than for pelvis signals. Moreover, the correlations between the NEVd-values are high and the deviations are generally small, rarely exceeding 21 mm for the head signals and 14 mm for pelvis.

We further provide histograms over the deviations for head and pelvis in Figure 7. The histograms include both *V* and *P*-values. In addition, these plots show the empirical estimates of normal distributions computed from the stride samples. The distributions of the absolute values for stride mean residuals are presented in Figure 8.

Finally, we provide mean RMSD values in Table 3 to give an estimate of curve similarity. Note that RMSD, as computed in Table 3, is sensitive to small time shifts. As the synchronization between signals $y_{sc}(t)$ and $y_{mc}(t)$ is approximate, the RMSD does not only reflect curve similarity but also errors in the synchronization.
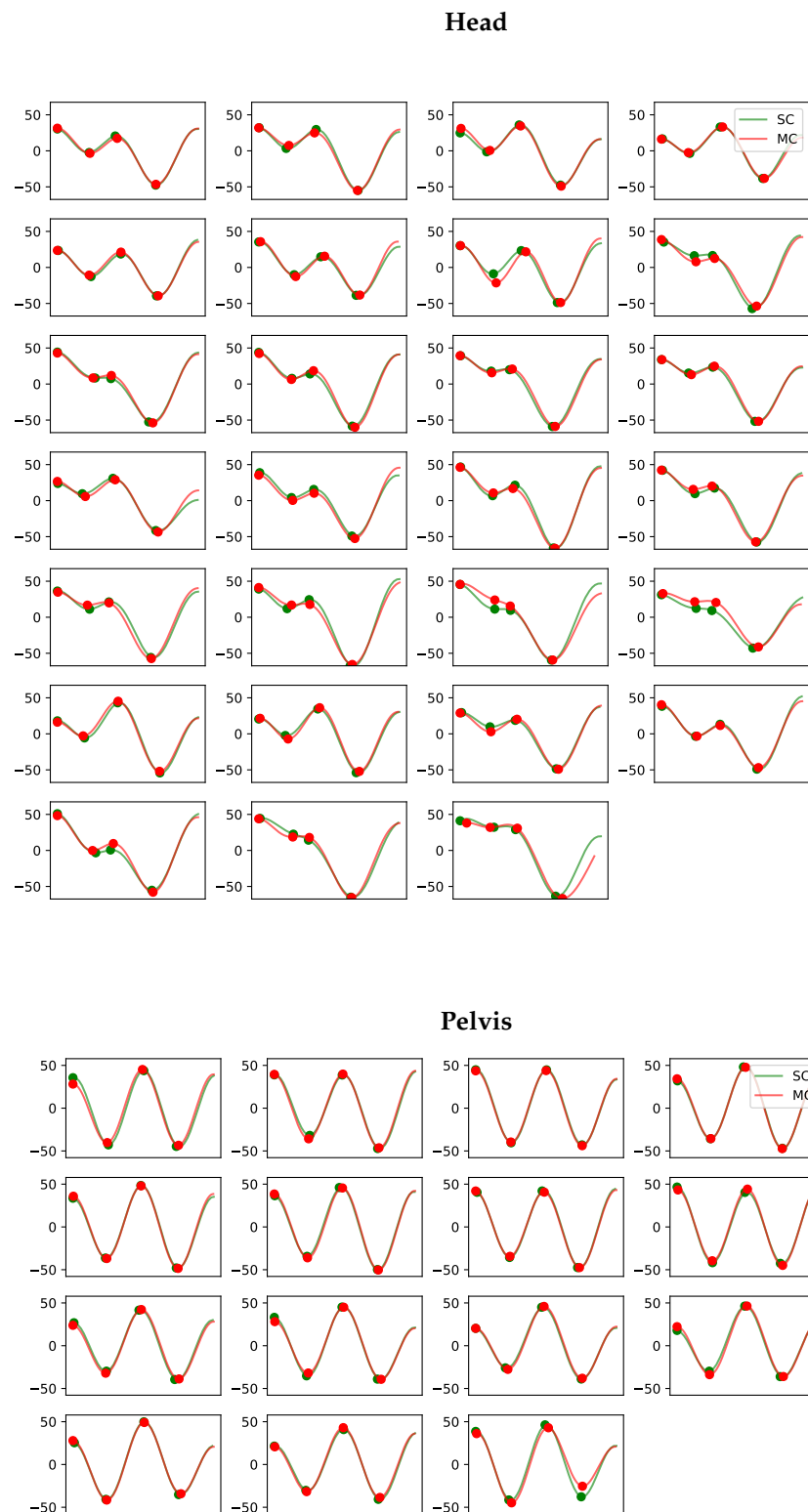
Head

Pelvis

Figure 4. Example of the vertical displacement signal per stride for the head and pelvis for horse 11. Each subplot contains the matched stride segments for the markerless single-camera system (SC) in green and the multi-camera marker-based motion capture system (MC) in red. The y-axis shows the vertical displacement in millimetres. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2. We observe that despite the high variability in curve shape between strides, there is a notable resemblance between the two systems.

Head *V*

Head *P*
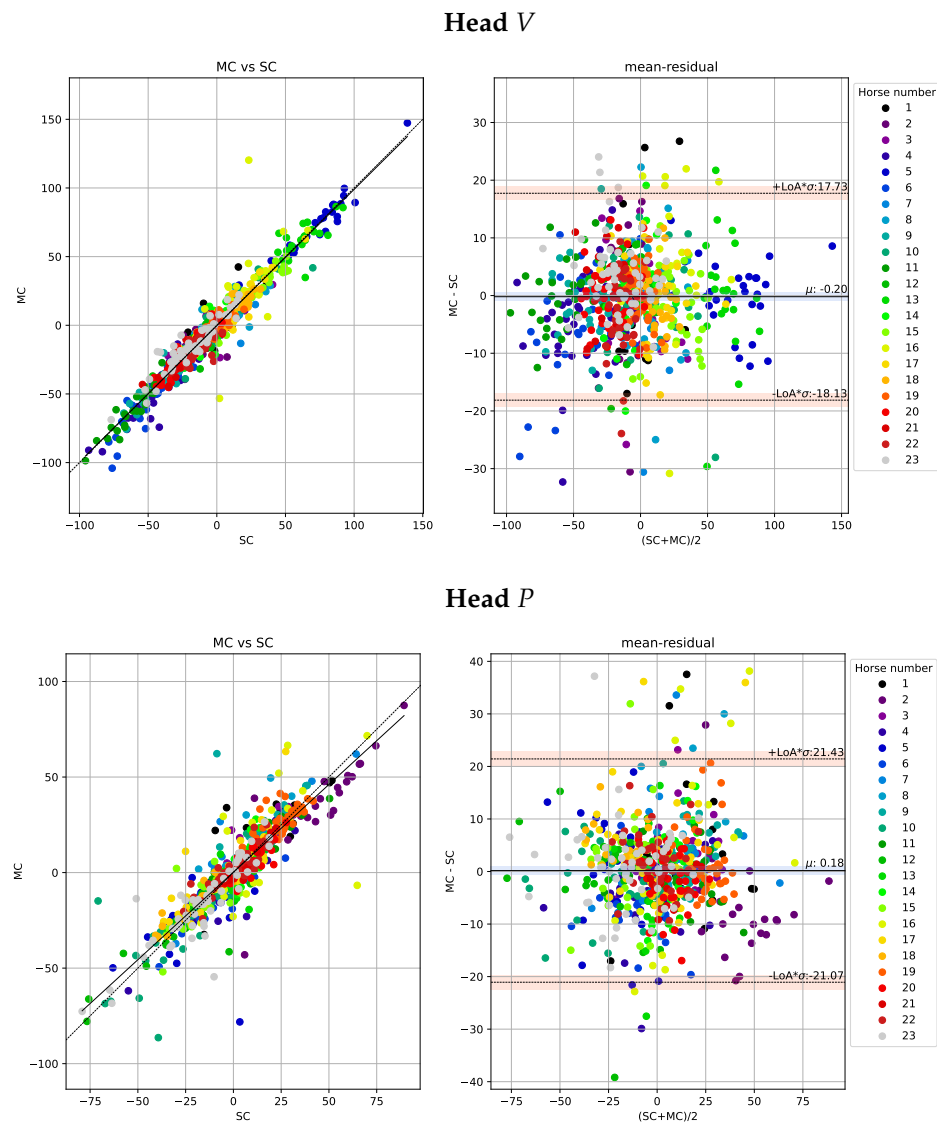
Figure 5. Scatter plots of the head metrics obtained per stride (n = 655) observed by the multi-camera marker-based motion capture system (MC) plotted against the observation by the single-camera markerless system (SC) are shown in the left sub-panels. Agreements between the systems, with limits of agreement (LoA) displayed as orange horizontal lines, are presented in the Bland-Altman plots in the right sub-panels. In the top row, we show the ΔV (MinDiff) and in the bottom row, we show ΔP (MaxDiff) defined in Section 2.7.1. For the purpose of visibility, we have fixed the range of the y-axis, causing a few samples with large residuals to be out of range. In the left plot, we have out of range residuals at 97.0 and −55.2, and in the right plot out of range residuals are at −48.9, −81.5, 70.9, 56.0, −47.1 −71.5, −44.4. While these samples have large errors they have little impact on the overall statistics.
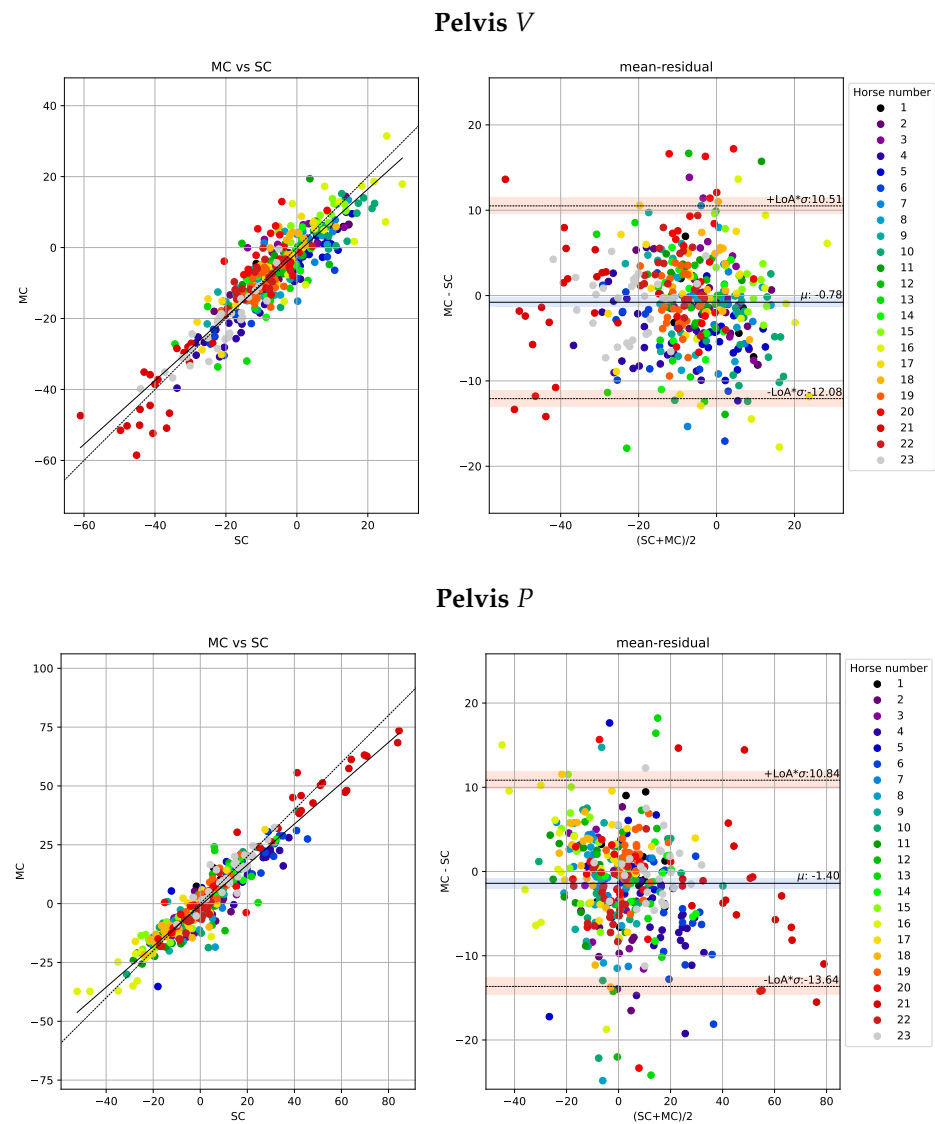
Figure 6. Scatter plots of the pelvic metrics obtained per stride (n = 404) observed by the multi-camera marker-based motion capture system (MC) plotted against the observation by the single-camera markerless system (SC) are shown in the left sub-panels. Agreements between the systems with limits of agreement (LoA) displayed as orange horizontal lines are presented in the Bland-Altman plots (right sub-panels). In the top row, we show the ΔV (MinDiff) and the bottom row we show ΔP (MaxDiff) defined in Section 2.7.1.
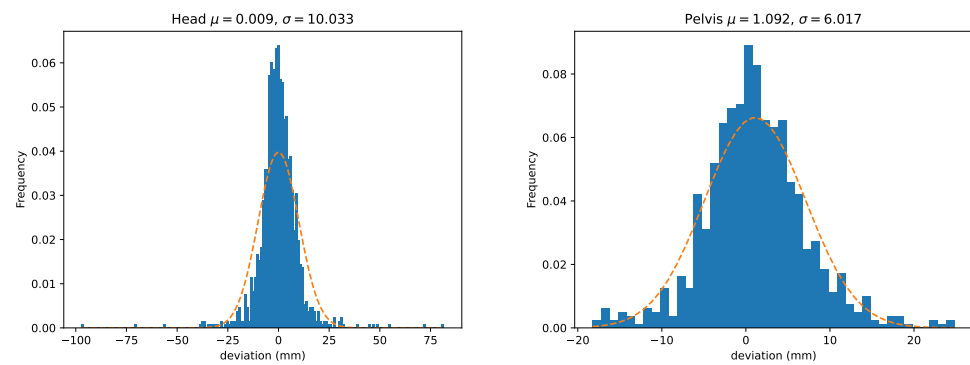
Figure 7. Distributions of the per-stride deviations for head (**left**) and pelvis (**right**). Here we have combined both peak and valley deviations defined in Section 2.7.1 (ΔV and ΔP). The dashed line displays the normal distribution estimated from the means and standard deviations of the between-systems deviations.
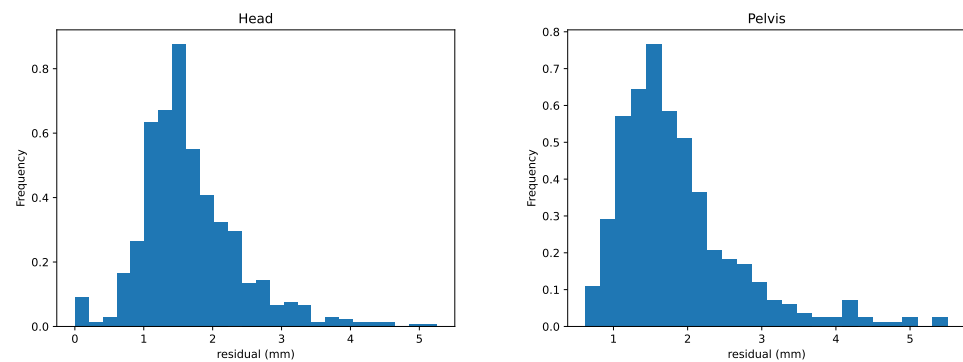
Figure 8. Distributions of the absolute values for stride mean residuals for head (**left**) and pelvis (**right**).

Table 3. Results summary for the measurement deviations between the systems over the entire dataset. In the top 6 rows we provide combined per-trial deviations ($\bar{D}$) using both deviations for the normalised MinDiff ($\Delta \bar{V}$) and MaxDiff ($\Delta \bar{P}$). For both head and pelvis, we compute the overall absolute mean, maximum and minimum deviations. In the bottom two rows we provide mean root mean square deviations (RMSD) as a comparison of the shape of the vertical displacement signals.

Per Trial	
head $\bar{D}$	2.2 mm
pelvis $\bar{D}$	2.2 mm
head max D	8.7 mm
pelvis max D	6.5 mm
head min D	0.0 mm
pelvis min D	0.0 mm
Per Stride	
head mean RMSD	5.0 mm
pelvis mean RMSD	3.5 mm

3.2. Per-Trial Comparisons

In this section, we provide the results from the per-trial comparison between MC and SC. In Table 3 we present statistics over the entire dataset from the per-trial mean NEVd-values. In this case, we combine $\bar{V}$ and $\bar{P}$. Thus, each stride contributes with two deviation

values. For these, we compute the overall absolute mean, maximum and minimum absolute deviations as,

$$\bar{D} = \frac{\sum_{m=1}^{M} |\Delta \bar{V}_m| + \sum_{m=1}^{M} |\Delta \bar{P}_m|}{2M} \tag{11}$$

$$\max D = \max(\{|\Delta \bar{V}_m|\}_{m=1}^{M} \cup \{|\Delta \bar{P}_m|\}_{m=1}^{M}) \tag{12}$$

$$\min D = \min(\{|\Delta \bar{V}_m|\}_{m=1}^{M} \cup \{|\Delta \bar{P}_m|\}_{m=1}^{M}), \tag{13}$$

where $M = 23$ is the number of trials, i.e number of horses in the dataset, and $\cup$ denotes the union of the sets of $\{|\Delta \bar{V}_m|\}_{m=1}^{M}$ and $\{|\Delta \bar{P}_m|\}_{m=1}^{M}$.

Similar to the per-stride comparison in Section 3.1, we use Bland-Altman plots [37] to inspect the statistical agreements. The plots for head and pelvis are shown in Figures 9 and 10, respectively. Not unexpectedly, the deviations between systems are smaller for the per-trial mean NEVd-values than for the per-stride comparison in Section 3.1, rarely exceeding 6.4 ms. Similar to the per-stride comparison, the differences between $\bar{V}$ and $\bar{P}$ are small. However, the per-trial deviations show similar values for pelvis and head, which could be due to the fact that the two systems jointly observed more strides for the head signal than the pelvis.
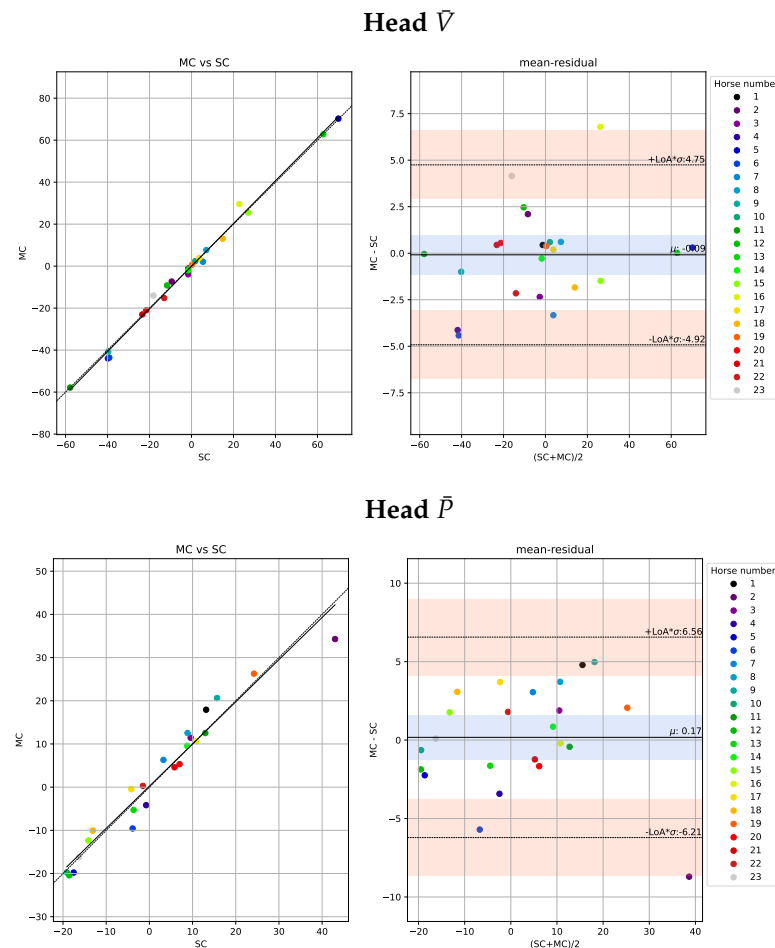
Figure 9. Trial-level scatter plots of the head metrics from the 23 included horses measured by the marker-based motion capture system (MC) versus the single-camera markerless system (SC) are presented in the left panel. Agreements between the systems, with limits of agreement (LoA) displayed as orange horizontal lines, are presented in the Bland-Altman plots in the right sub-panels. In the top row, we show $\Delta \bar{V}$ (trial mean MinDiff) and in the bottom row, we show $\Delta \bar{P}$ (trial mean MaxDiff) defined in Section 2.7.1.

In addition, we provide histograms over all the deviations (both $\bar{V}$ and $\bar{P}$) for head and pelvis in Figure 11. These plots also show the empirical estimates of normal distributions computed from the stride samples.
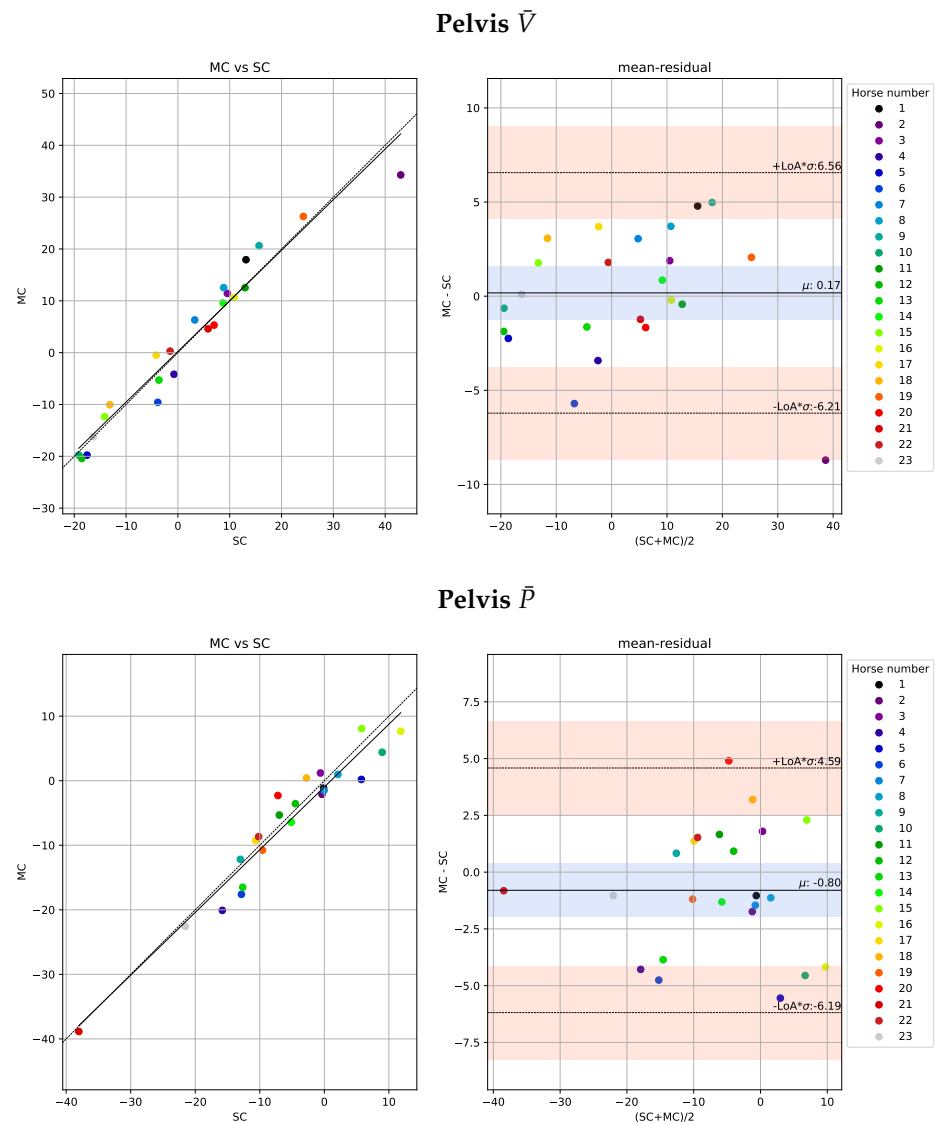
Figure 10. Trial-level scatter plots of the pelvic metrics from the 23 included horses measured by the marker-based motion capture system (MC) versus the single-camera markerless system (SC) are presented in the left panel. Agreements between the systems, with limits of agreement (LoA) displayed as orange horizontal lines, are presented in the Bland-Altman plots in the right sub-panels. In the top, we show $\Delta\bar{V}$ (trial mean MinDiff) and in the bottom row, we show $\Delta\bar{P}$ (trial mean MaxDiff) defined in Section 2.7.1.
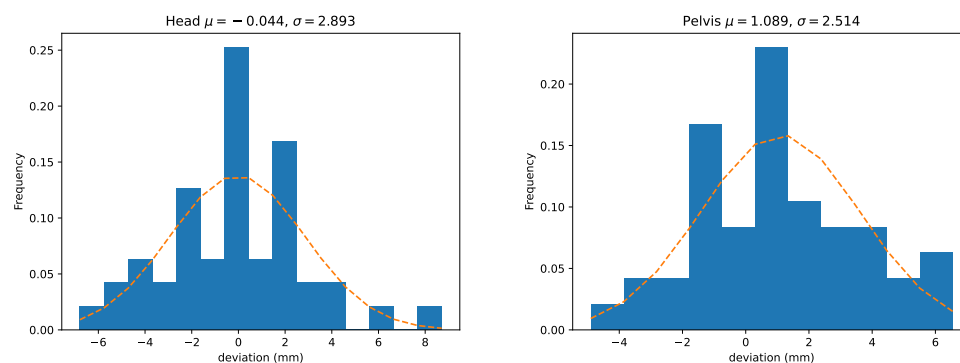
Figure 11. Distributions of the per-trial deviations for head (**left**) and pelvis (**right**). Here we have combined both peak and valley deviations defined in Section 2.7.1 ($\Delta \bar{V}$ and $\Delta \bar{P}$). The dashed line shows the normal distribution estimated from the means and standard deviations of the deviations.

4. Discussion

In this work, we demonstrated that deep neural networks and computer vision can be applied to reliably perform orthopaedic gait analysis for horses when trotted in hand on the straight during lameness assessment. The recorded average per-trial errors of 2.17 mm (head) and 2.19 mm (pelvis) are well below the previously recorded between-measurement-variation, of 18 mm (head) and 6 mm (pelvis) of horses trotting in a MC-system with several repetitions over two consecutive days and a one-month follow-up [21]. The results thus indicate that the average per-trial errors of the SC-extracted variables compared to the MC-extracted variables were small enough to not be a major hindrance when used for objective lameness assessment.

From a clinical perspective, the ease-of-use of the SC-system studied has a clear benefit since it allows affordable, repeated observations of equine patients, which can help to understand the considerable between trial variations observed in gait asymmetry [21].

When comparing the results of this study to another validation study performed on an IMU-based system (compared to a MC-system), it was found that stride-by-stride limits of agreement for the pelvic variables were approximately twice the magnitude for the SC-system described in this work [20]. It has to be acknowledged that comparing results from different samples can be a confounding factor in this case, but it still presents a general indication of how a computer vision based solution may compare to an IMU-based solution. Unfortunately, limits of agreement for the IMU-system in the Bosch et al. study [20] versus our SC-system can not be compared for the head variables since these were not presented in the IMU-system validation study.

The reported accuracy of the SC-system could also be compared to a previous test-retest repeatability study of an IMU-system, where the 95% confidence interval was reported as approximately 6 mm for head asymmetries and 3 mm for pelvic asymmetries [38]. However, a later study comparing that IMU-system to a different IMU-system, also developed for detecting equine lameness, found that the limits of agreement between the two systems were in this same range [36]. Further, it was found that the system used by [38] consistently underestimated the amount of movement asymmetry compared to the other IMU-system, which had previously been shown to give values comparable to optical motion capture. It has been suggested that the confidence intervals reported by Keegan et al [38] should be adjusted to 8 and 4 mm [13] for the other IMU-system, and this is likely a suitable adjustment also for an MC-system.

There were two outlier trials with a recorded per-trial mean deviation of 8.71 mm (head) and 6.54 mm (pelvis). Notably, these outliers occurred for horses with large asymmetries (see horse 2 in Table 1 and horse 4 in Table 2), and did not change the sign of the calculated variables or the clinical interpretation of the gait data. Hence, these deviations would not confuse which limb was affected by the asymmetry. We hypothesized that these errors might be due to difficulties in detecting and tracking the head and pelvis when they

were occluded e.g., the horse lowering its head and hiding it behind the trunk, or lifting the tail obscuring the pelvic region.

Another SC-system utilizing deep neural networks to detect and track trotting horses for the purpose of lameness assessment has been previously described [39]. Although their approach was similar to the method presented in this study, the authors did not quantify movement symmetry metrics directly. Instead, they focused on lame limb classification. It is also worthwhile noting that a SC-system has been developed and shown to be able to perform reliable gait analysis of human subjects [27]. The level of precision described is said to open up for several potential applications in human medicine.

Is markerless more or less? A dichotomous answer cannot be given here. There are technical drawbacks related to using computer vision techniques implemented in a SC-system aimed towards objective equine lameness assessment, lower accuracy and lack of 3D motion to name two. However, these drawbacks have to be weighed against the benefit of having a lightweight, portable and low-cost system available for data collection, that allows repeatable observations of the horse. Other systems, such as IMUs and MC are typically more expensive, are sometimes limited to laboratory environments and require more interaction in terms of placing markers and sensors on the horse. Inevitably, this leads to fewer measurements being done and it is well known that low sample sizes of horses are the standard in many equine biomechanics studies. This study has shown that a simple application on a smartphone can be a tool for flexible and reliable collection of kinematic asymmetry data from horses. By extension, this opens up opportunities for larger-scale biomechanics research studies in non-laboratory environments. Coupling this with current advances in machine learning, where computers efficiently learn from data to perform predictions, we suggest that SC can be used to accelerate our understanding of horse locomotion and horse welfare.

Limitations

In the current study, all measurements were performed in the same clinical indoor environment on a limited number of horses (n = 23). The SC-system would likely be more challenged if there was a severe lack of light or if for example heavy rain obscured the visibility of the horse in the video. However, testing under such conditions was not within the scope of this study and would have been impossible to perform given that the state-of-the-art MC-system is a permanent indoor installation. The smartphone was placed on a tripod during the data collection, as handheld recordings would demand a stabilisation algorithm to be applied to the video. As such, further research is required to evaluate the SC-system under handheld conditions. Also, the length of the runway was 30 m, hence a greater distance between the camera and the horse has not been investigated. We did, however, analyse the error per stride index and did not find increasing deviations of the lameness metrics studied.

There are today several iPhone models with different camera specifications, such as image resolution and sampling rate. This study was limited to the use of an iPhone12 Pro Max where the resolution was set to 2160×3840 pixels and the frame rate to 60 Hz. Newer iPhone models come with even better camera specifications. Further research should investigate whether these models would improve the output from the SC-system even further.

The neural networks used in the SC-system were trained on image data of horses that did not have markers attached to the skin. Therefore, the skin-attached reflective markers used in this validation study could be suspected to partly obscure the anatomical region of interest and potentially present a problem to the SC-system. Visual inspection of the tracking from the SC-system, did however show very stable detection of the anatomical segments. This is confirmed by the comparison to the output of the MC-system.

This study only investigated vertical movement asymmetry of head and pelvis in a straight line. 3D motion comparisons would be of interest in the future, in order to provide a more detailed analysis of horse locomotion e.g., limb retraction and protraction

angles studied from lateral view or on a circle. However, 3D lifting from a 2D image is an inherently complicated computer vision task where more advanced methods would be required, and so these kinematic variables were not investigated in this preliminary study.

5. Conclusions

We conclude that objective gait analysis for lameness assessment in horses can be reliably performed using a smartphone and computer vision analysis built on deep neural networks. The measurement deviation, when compared to a state-of-the-art motion capture system, is larger compared to IMU-based systems [20], but the error is clearly lower than observed levels of "between trial variation" from earlier studies [21]. The ease-of-use of the system makes repeated observations of a horse's lameness more feasible, which can provide more objective data points for treatment evaluation.

Abbreviations

The following abbreviations are used in this manuscript:

SC	Single-camera markerless system
MC	Multi-camera marker-based system
VDS	Vertical displacement signal
NEVd	Normalised extreme value differences
MaxDiff	Difference between local maxima values within a stride
MinDiff	Difference between local minima within a stride
P	Normalised difference between local maxima values of the VDS per stride
V	Normalised difference between local minima values of the VDS per stride
$\bar{P}$	Trial mean P
$\bar{V}$	Trial mean V
σP	Trial standard deviation of P
σV	Trial standard deviation of V

ΔP	Deviation between two corresponding P's from different systems
ΔV	Deviation between two corresponding V's from different systems
$\Delta \bar{P}$	Deviation between two corresponding $\bar{P}$'s from different systems
$\Delta \bar{V}$	Deviation between two corresponding $\bar{V}$'s from different systems
$\bar{D}$	Mean absolute deviation over dataset
$\max D$	Maximum absolute deviation over dataset
$\min D$	Minimum absolute deviation over dataset
R	Range of motion of the VDS per stride
LDA	Linear Discriminant Analysis
LoA	Limit of Agreement

Appendix A. All Stride Curves
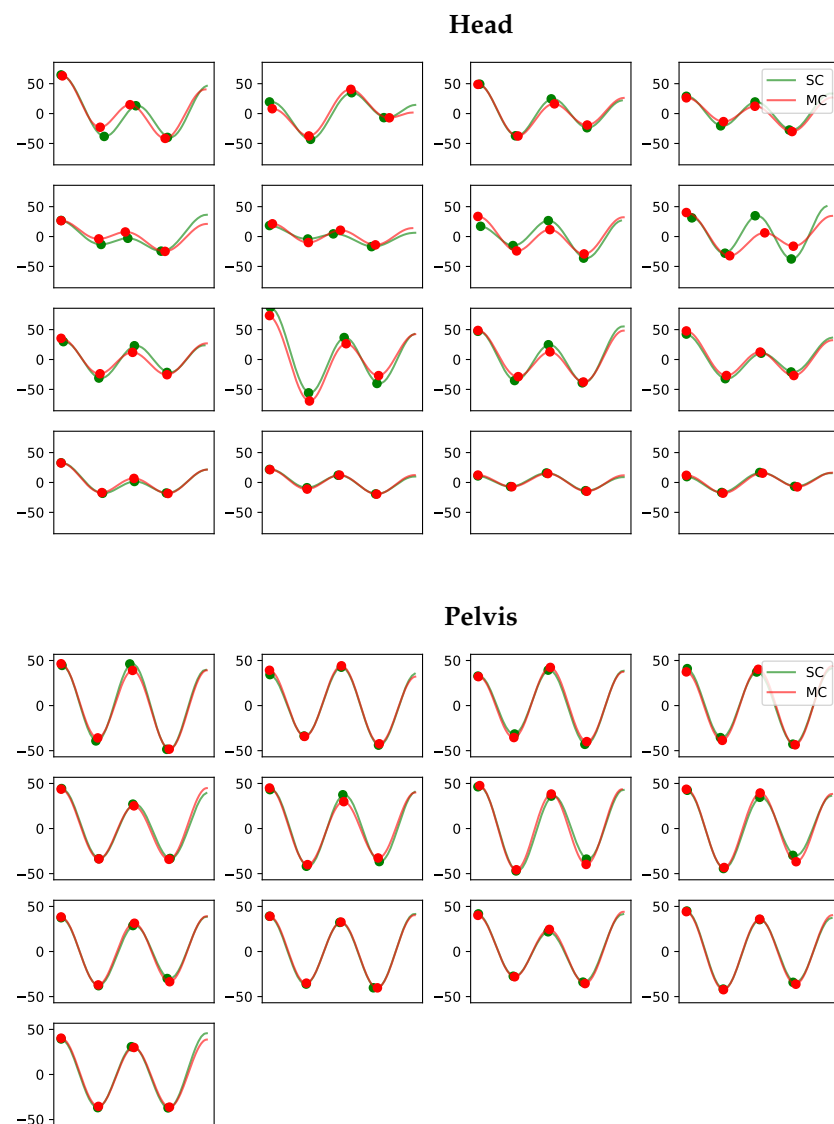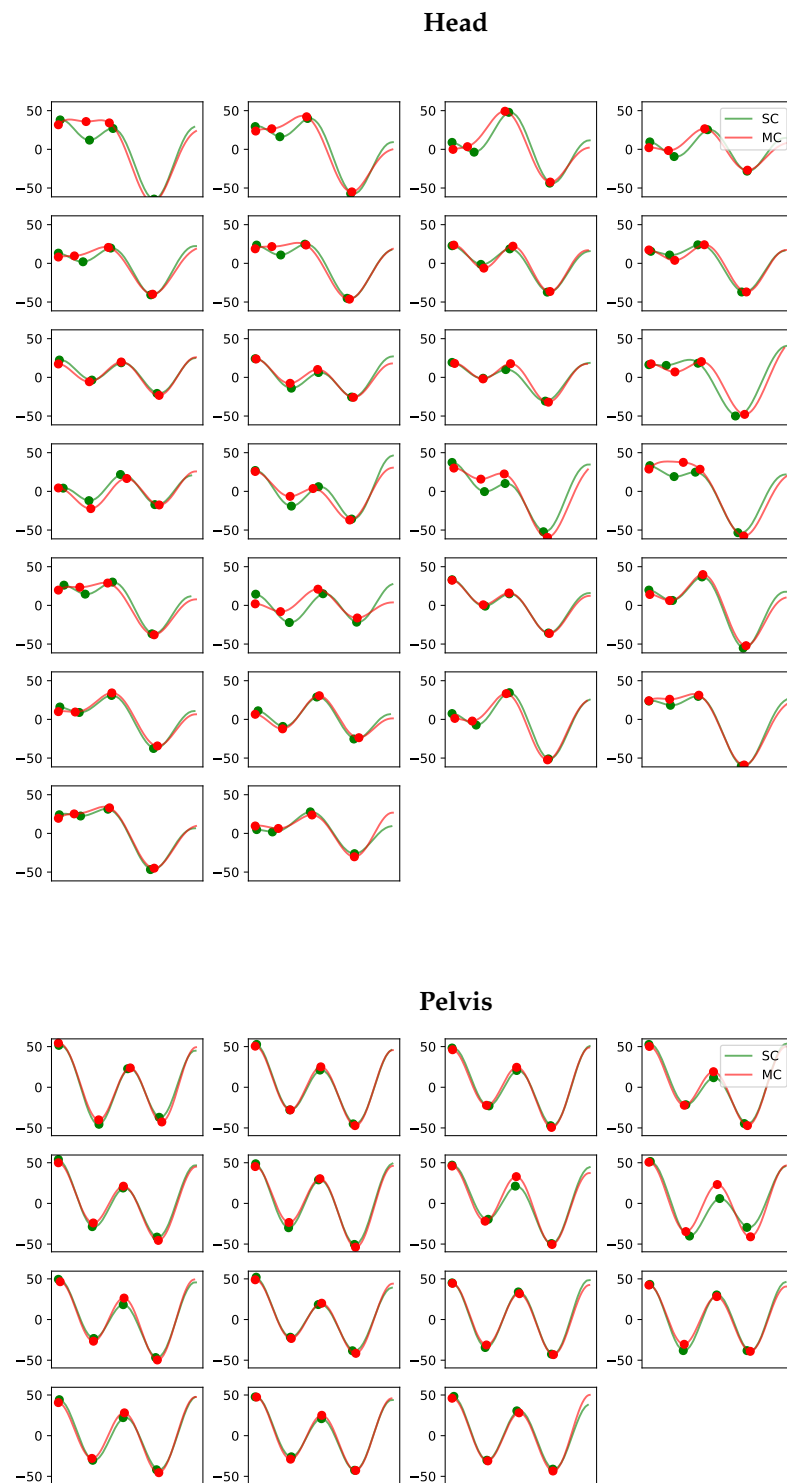
Head

Pelvis

Figure A1. Vertical displacement signals per stride for the head and pelvis for horse 1. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
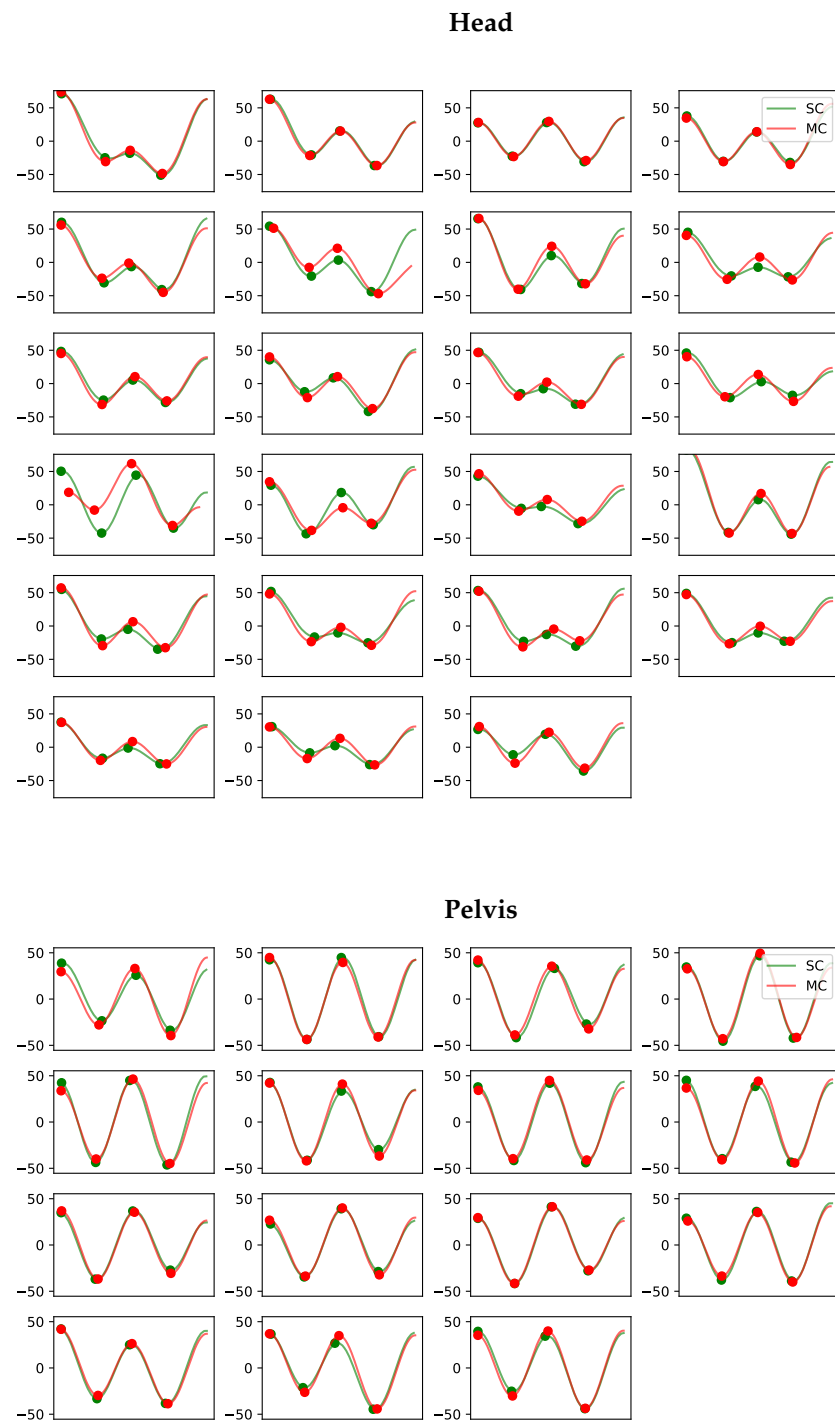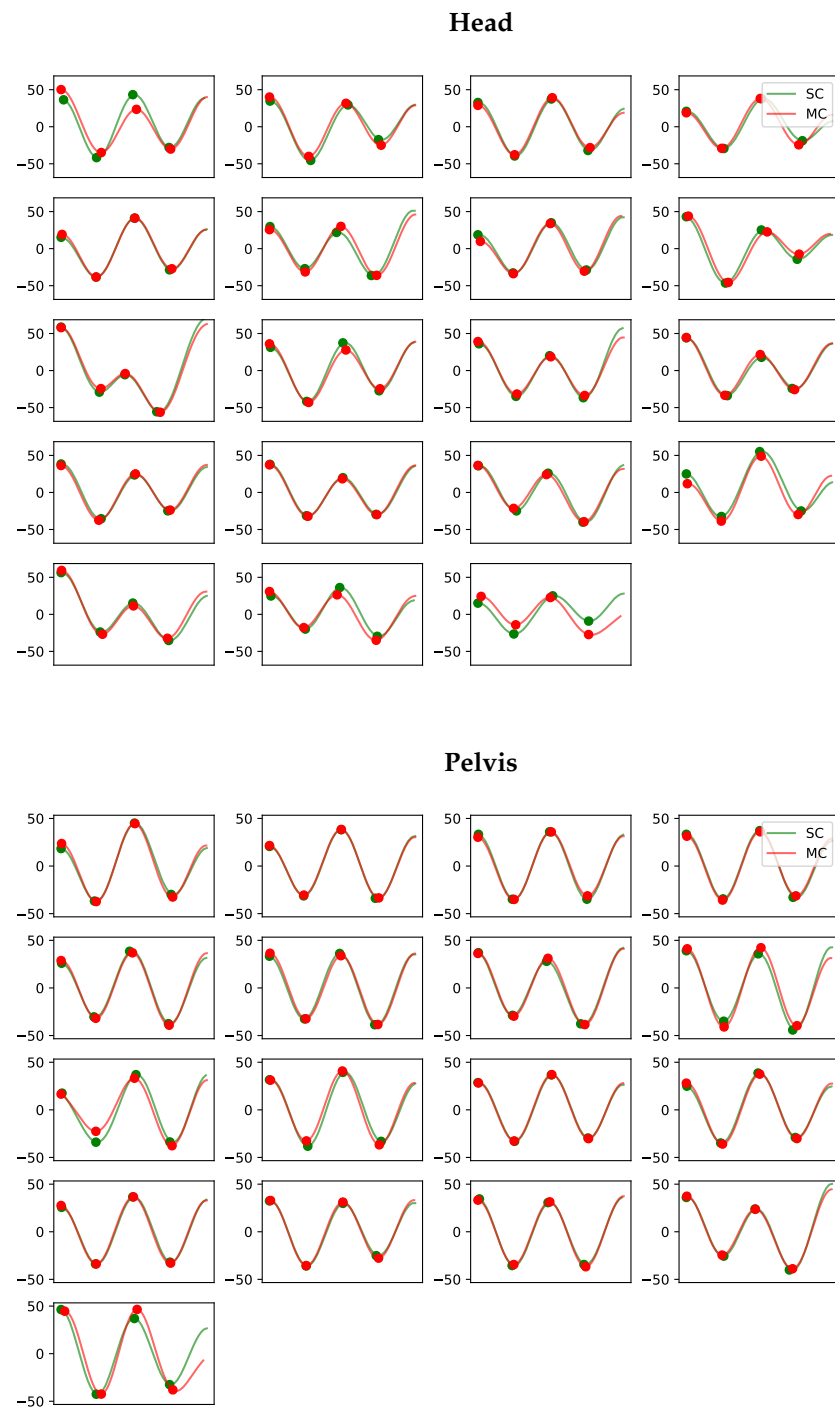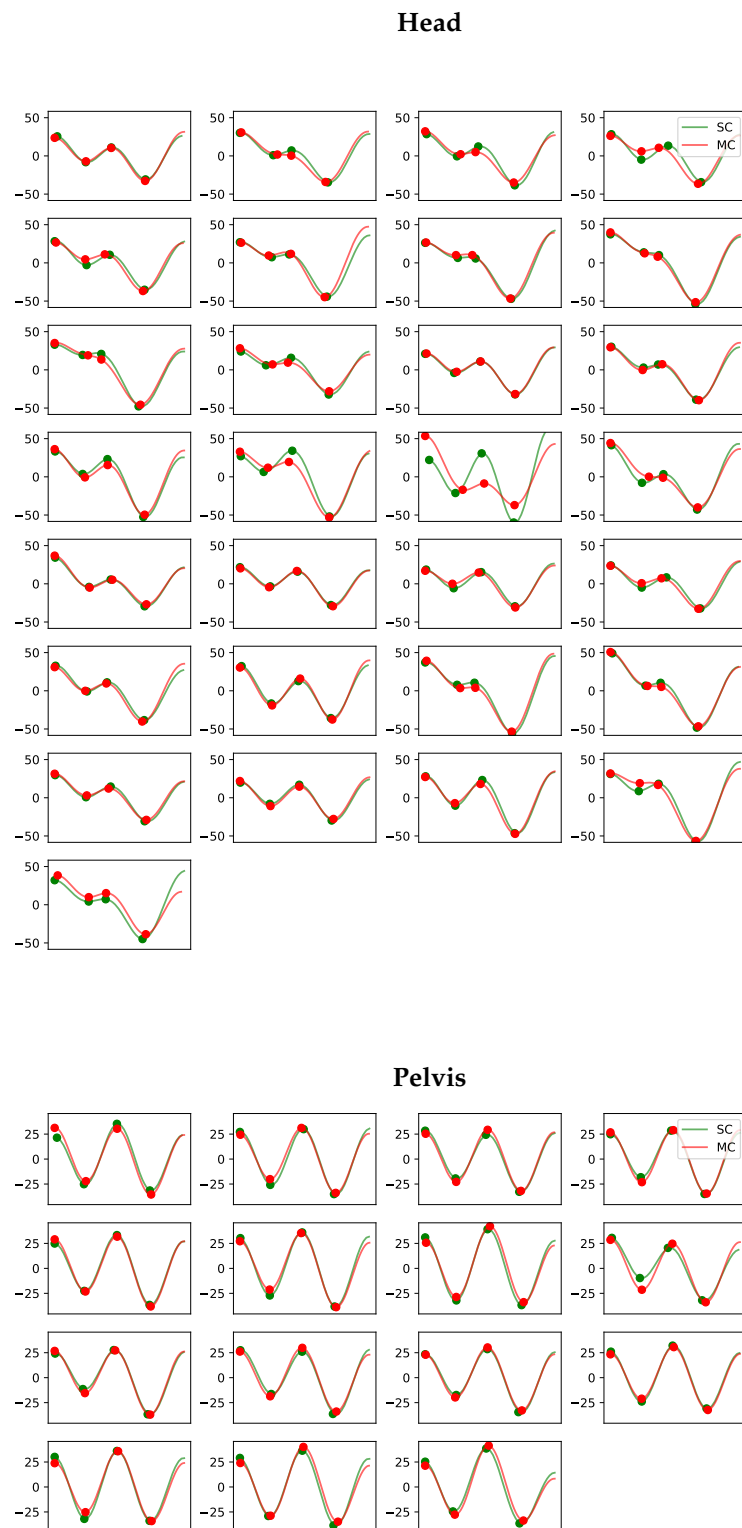
Head

Pelvis

Figure A2. Vertical displacement signals per stride for the head and pelvis for horse 2. Each subplot contains the matched stride segments the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
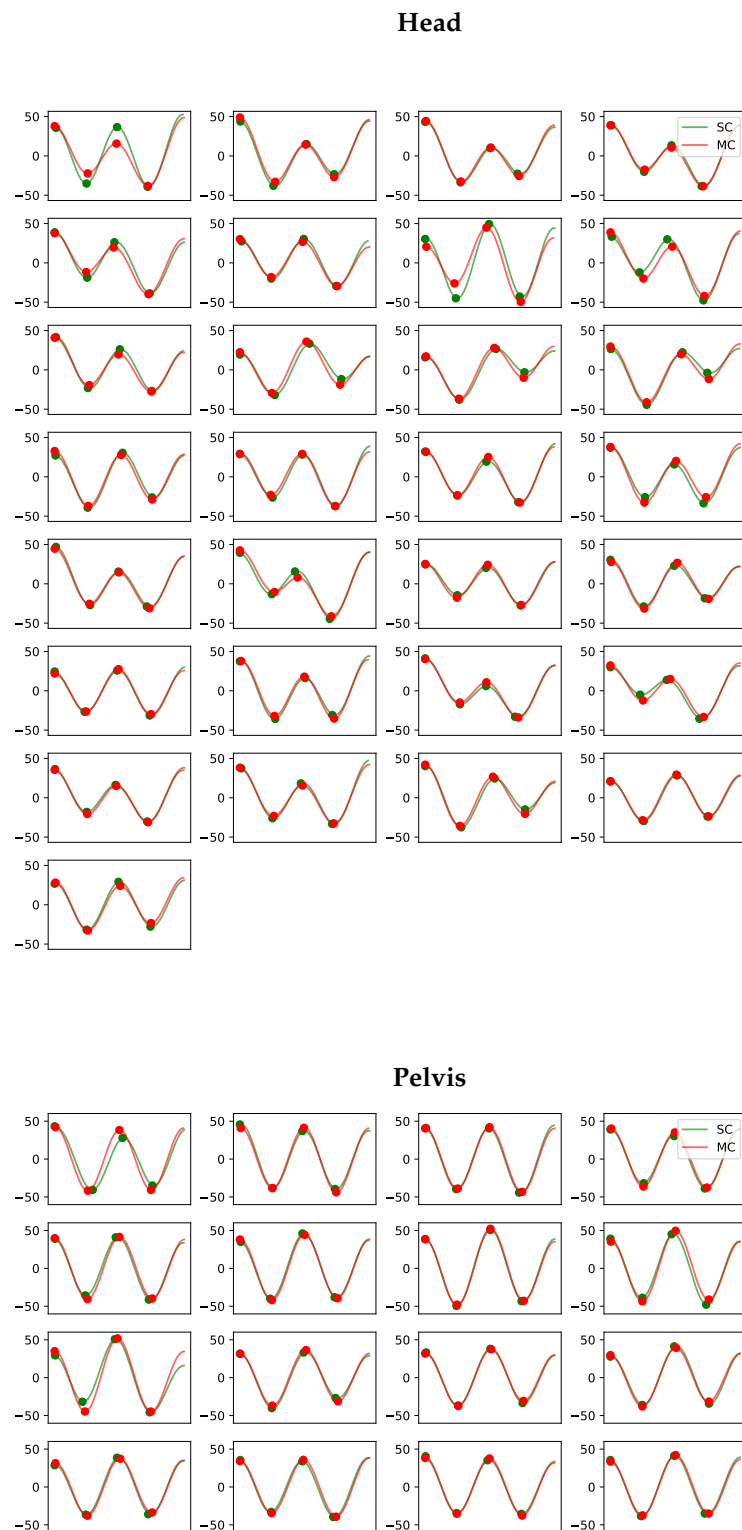
Head

Pelvis

Figure A3. Vertical displacement signals per stride for the head and pelvis for horse 3. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
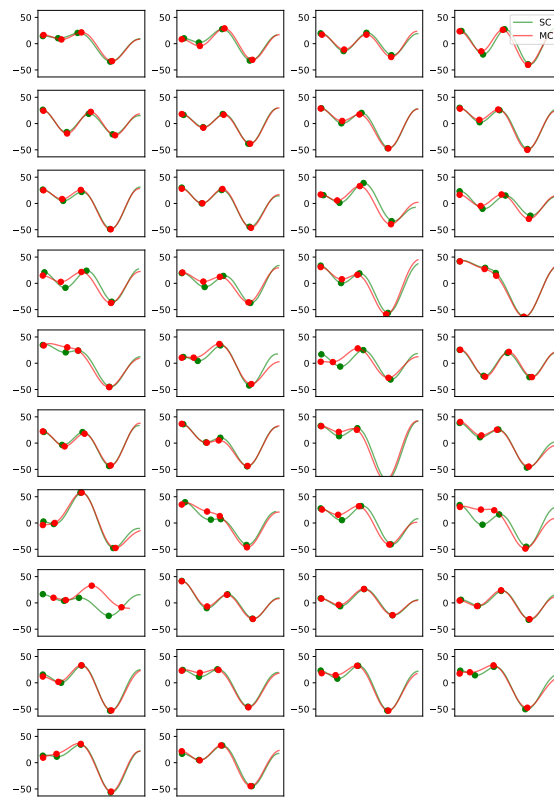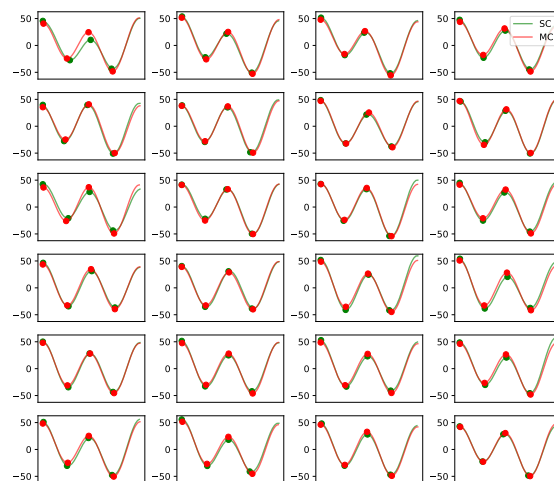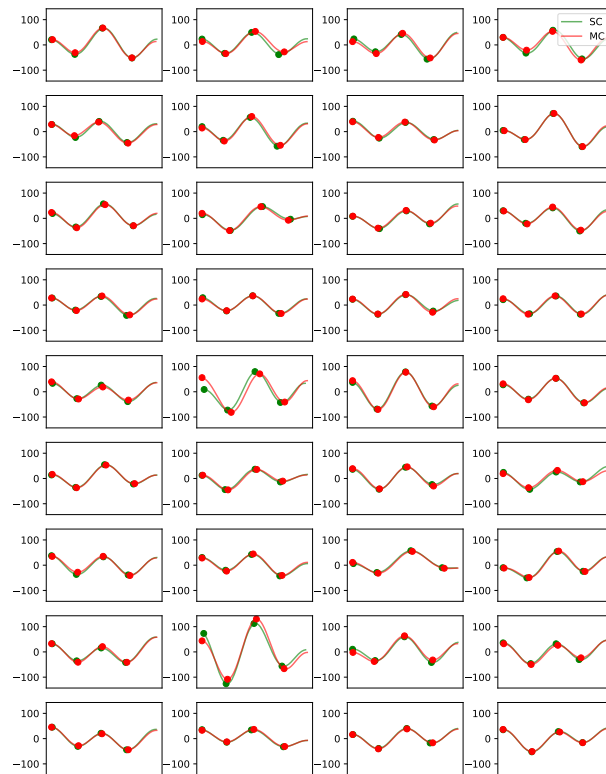
Head

Pelvis

Figure A4. Vertical displacement signals per stride for the head and pelvis for horse 4. Each subplot contains the matched stride segments forthe single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
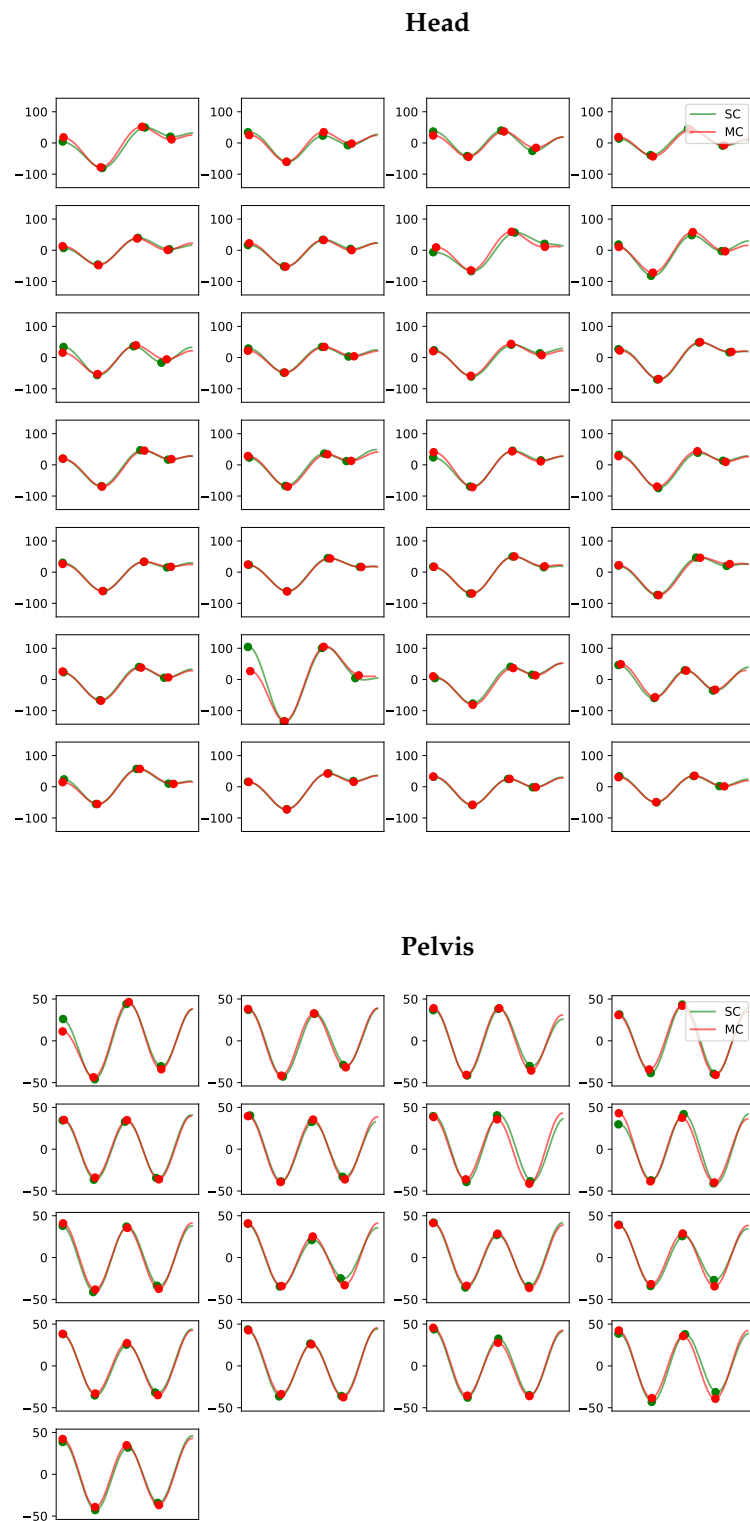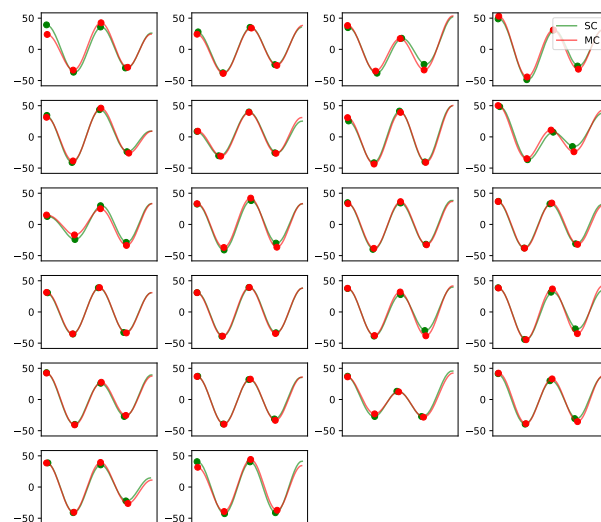
Head

Pelvis

Figure A5. Vertical displacement signals per stride for the head and pelvis for horse 5. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
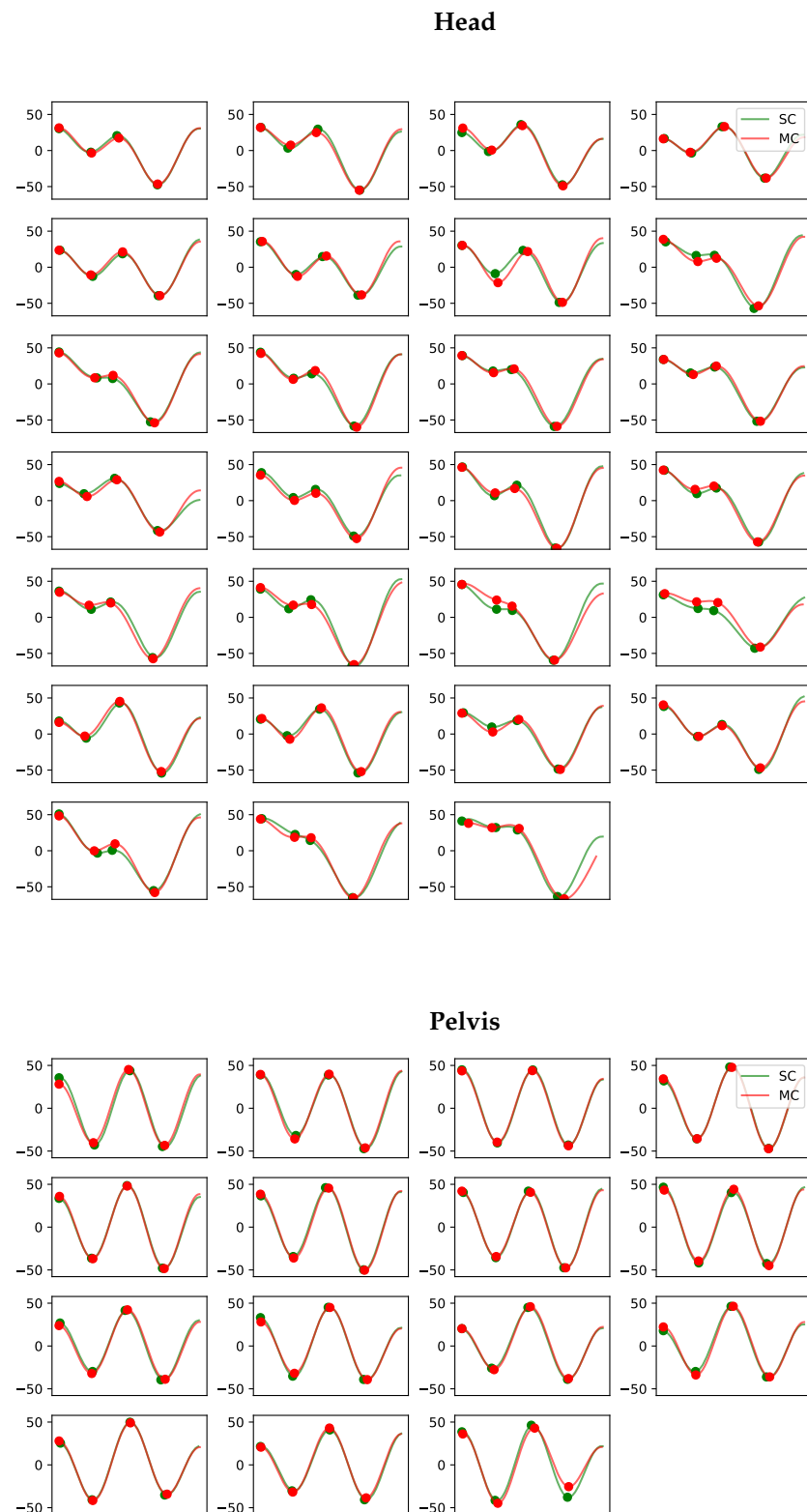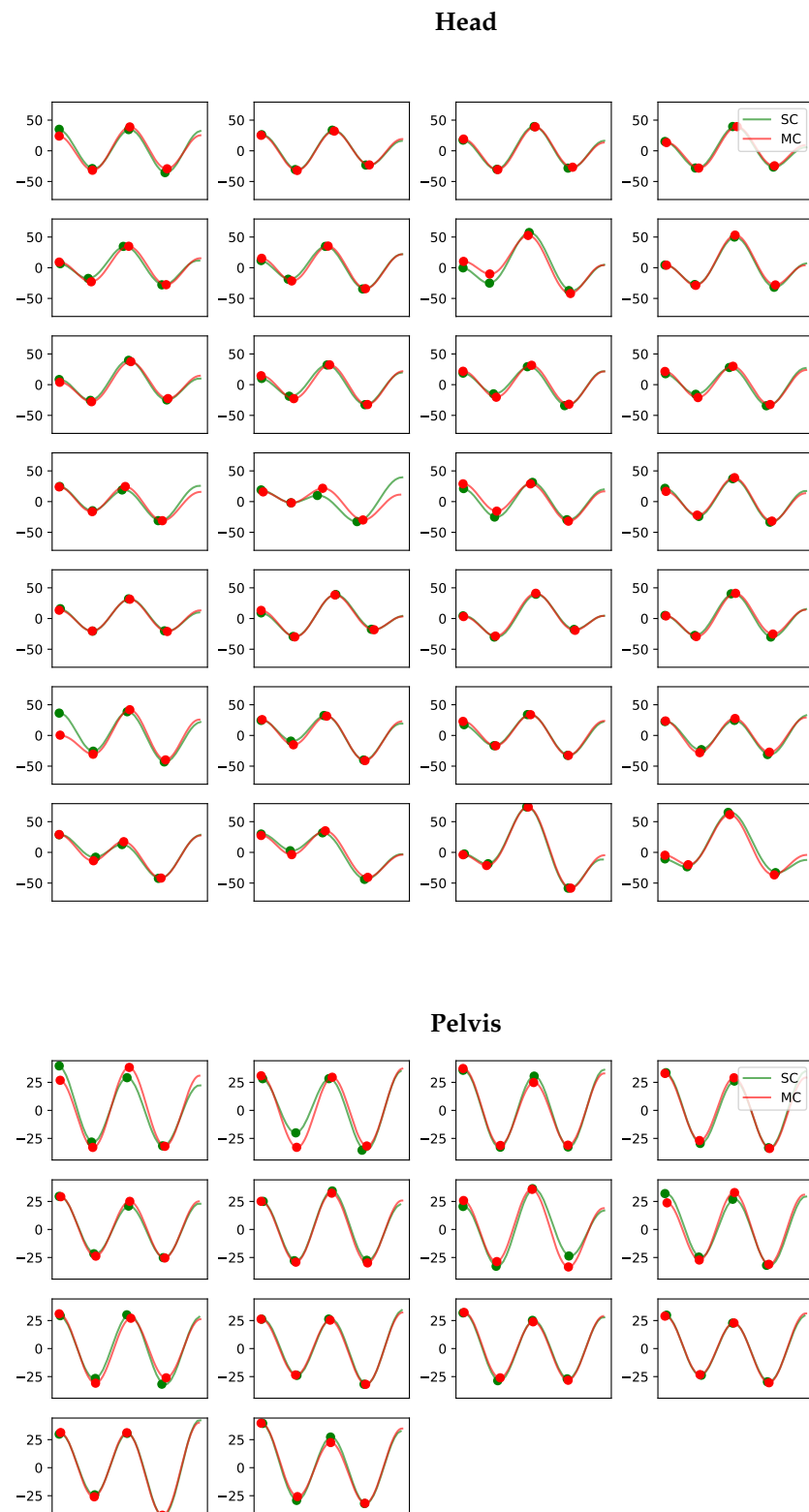
Head

Pelvis

Figure A6. Vertical displacement signals per stride for the head and pelvis for horse 6. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
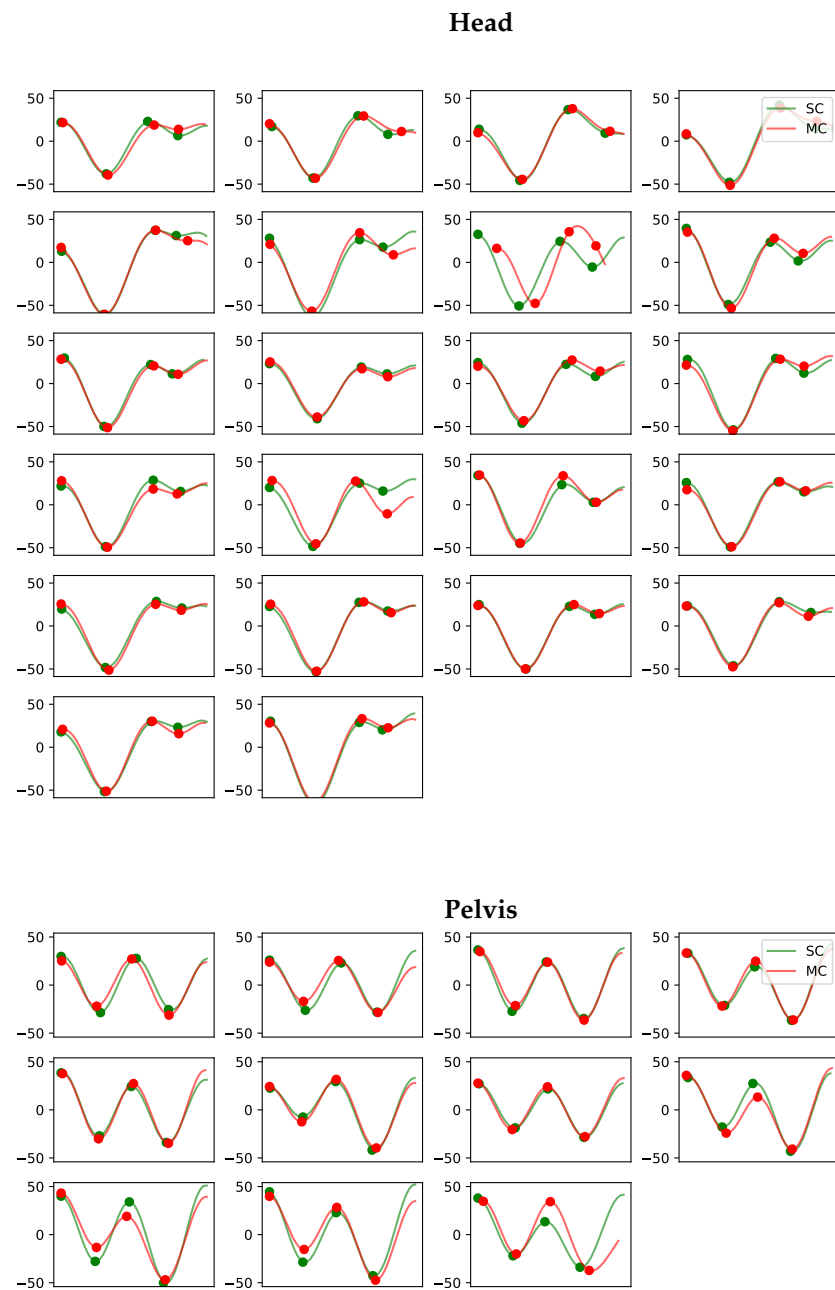
Head

Pelvis

Figure A7. Vertical displacement signals per stride for the head and pelvis for horse 7. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
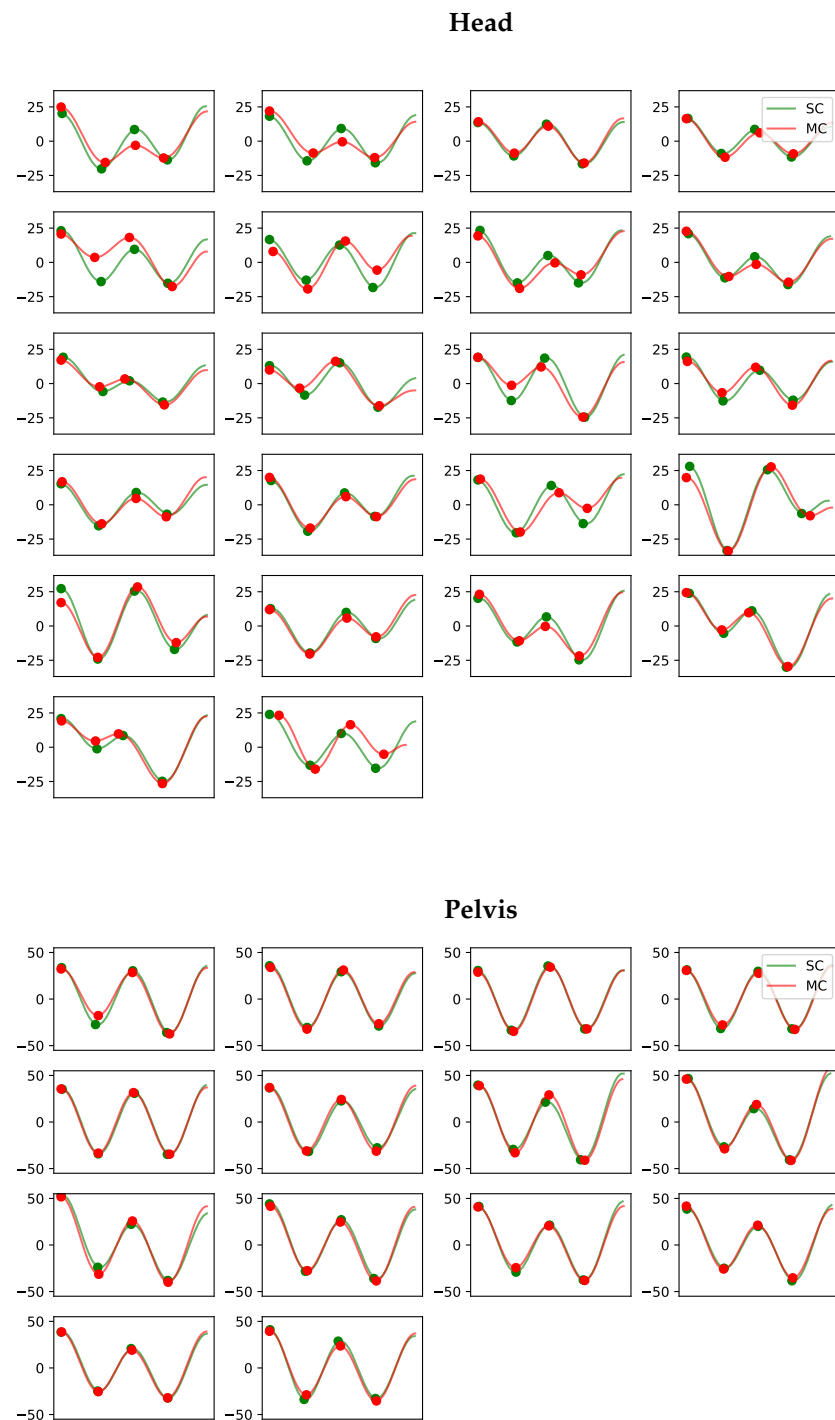
Head

Pelvis

Figure A8. Vertical displacement signals per stride for the head and pelvis for horse 9. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.

Head

Pelvis

Figure A9. Vertical displacement signals per stride for the head and pelvis for horse 10. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
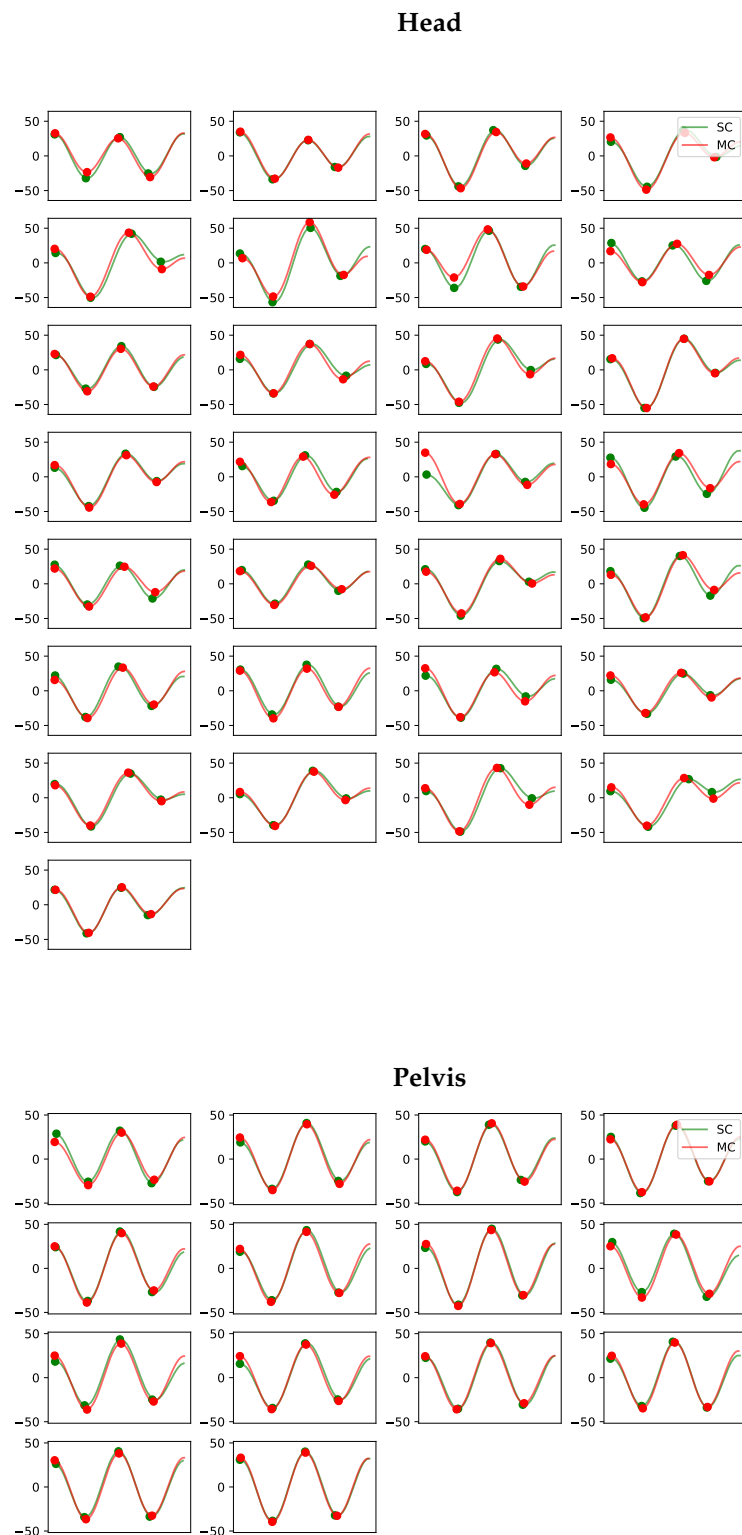
Head

Pelvis

Figure A10. Vertical displacement signals per stride for the head and pelvis for horse 11. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
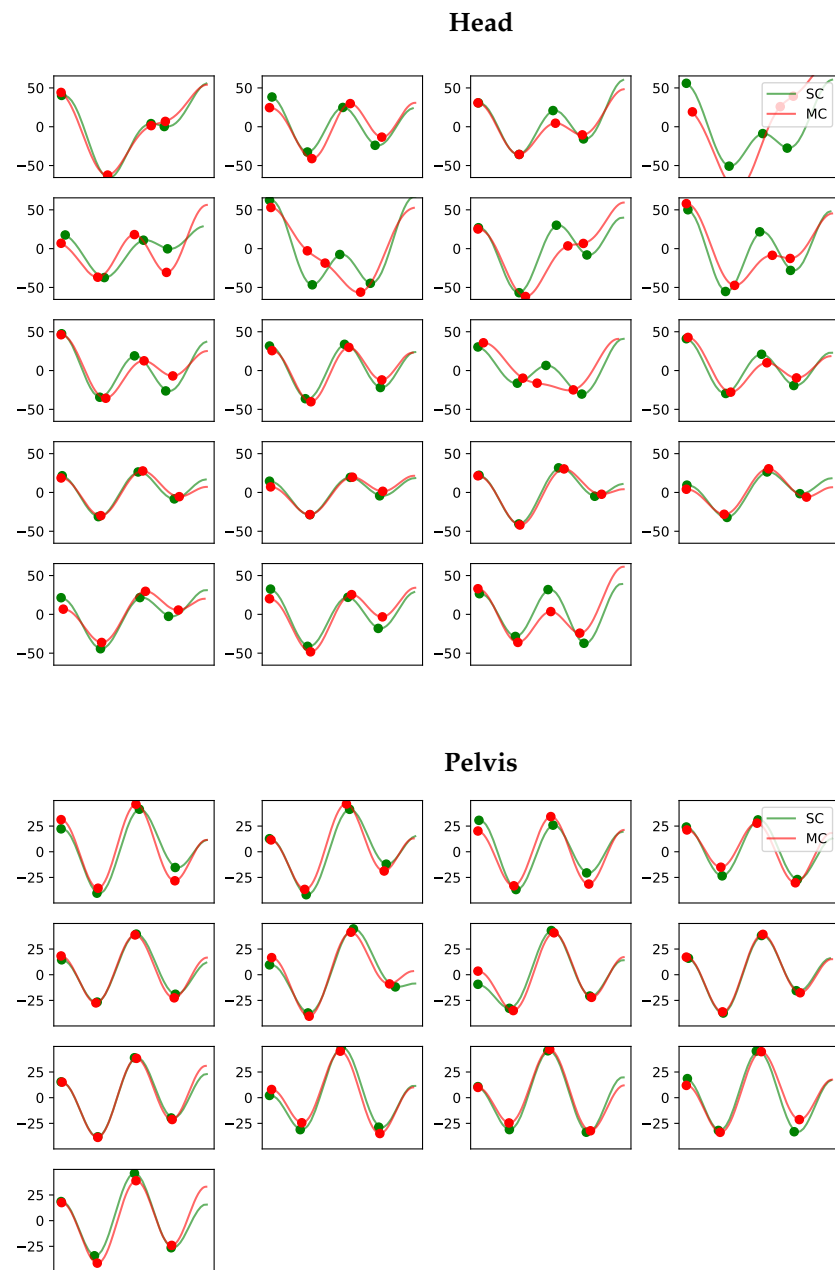
Head

Pelvis

Figure A11. Vertical displacement signals per stride for the head and pelvis for horse 12. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
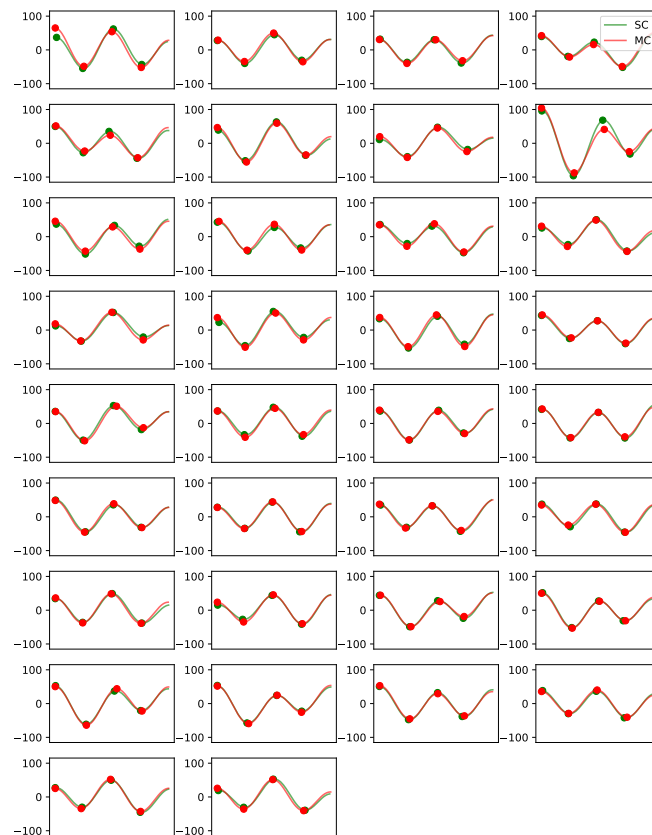
Head

Pelvis

Figure A12. Vertical displacement signals per stride for the head and pelvis for horse 13. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
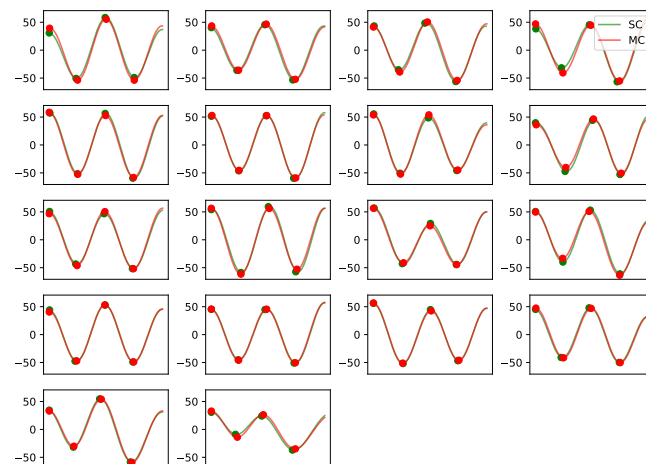
Head

Pelvis

Figure A13. Vertical displacement signals per stride for the head and pelvis for horse 14. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.

Head

Pelvis

Figure A14. Vertical displacement signals per stride for the head and pelvis for horse 15. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
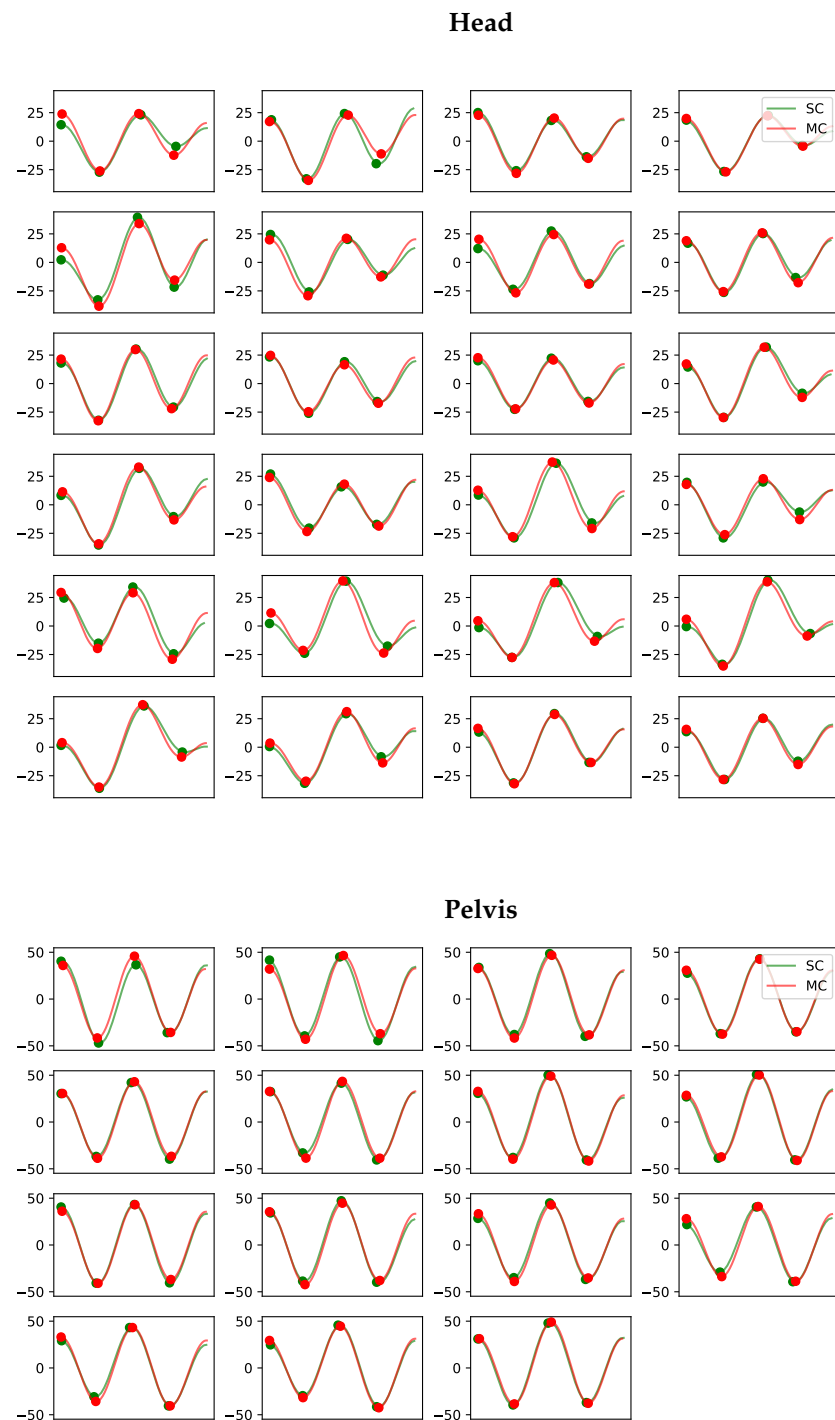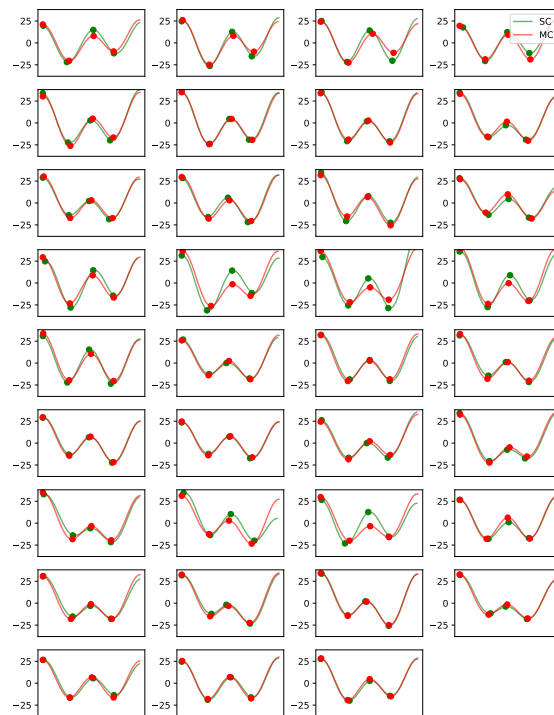
Head

Pelvis

Figure A15. Vertical displacement signals per stride for the head and pelvis for horse 16. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
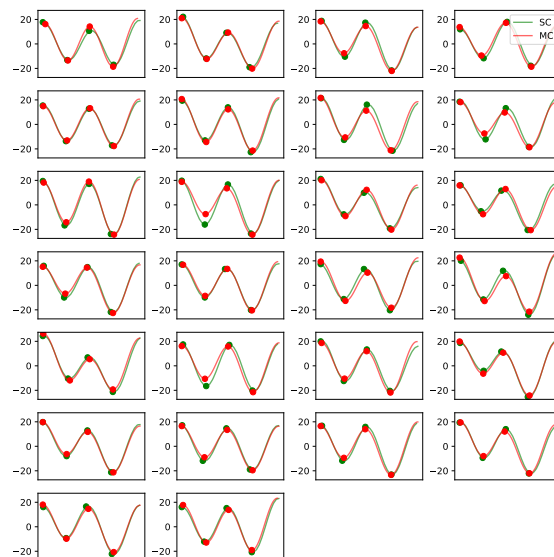
Head

Pelvis

Figure A16. Vertical displacement signals per stride for the head and pelvis for horse 17. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
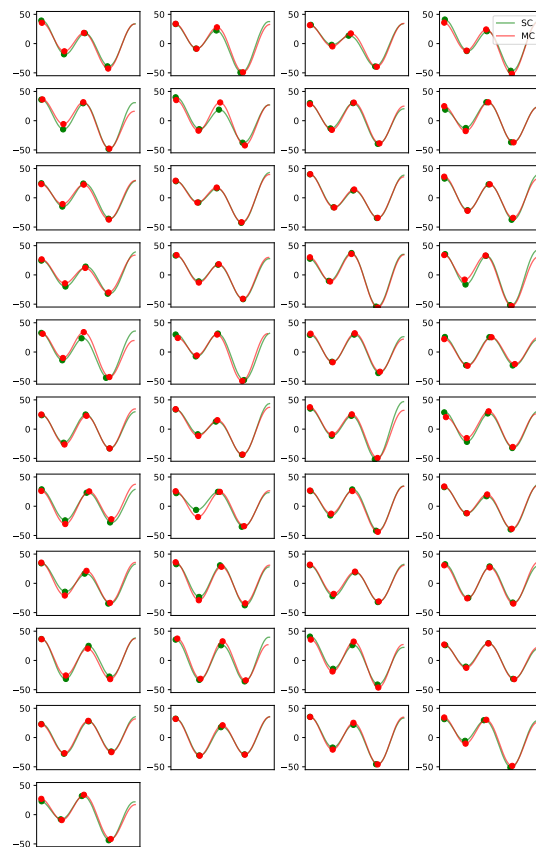
Head

Pelvis

Figure A17. Vertical displacement signals per stride for the head and pelvis for horse 18. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
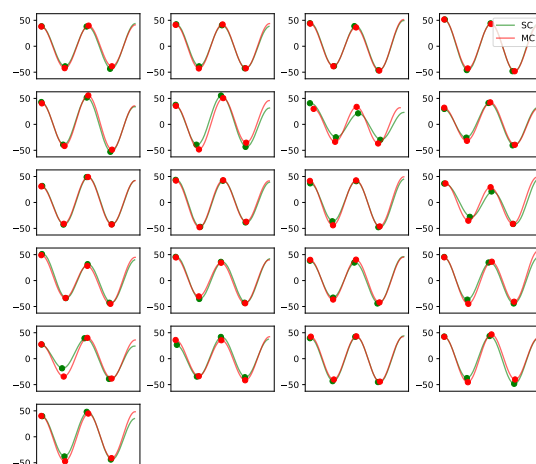
Head

Pelvis

Figure A18. Vertical displacement signals per stride for the head and pelvis for horse 19. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
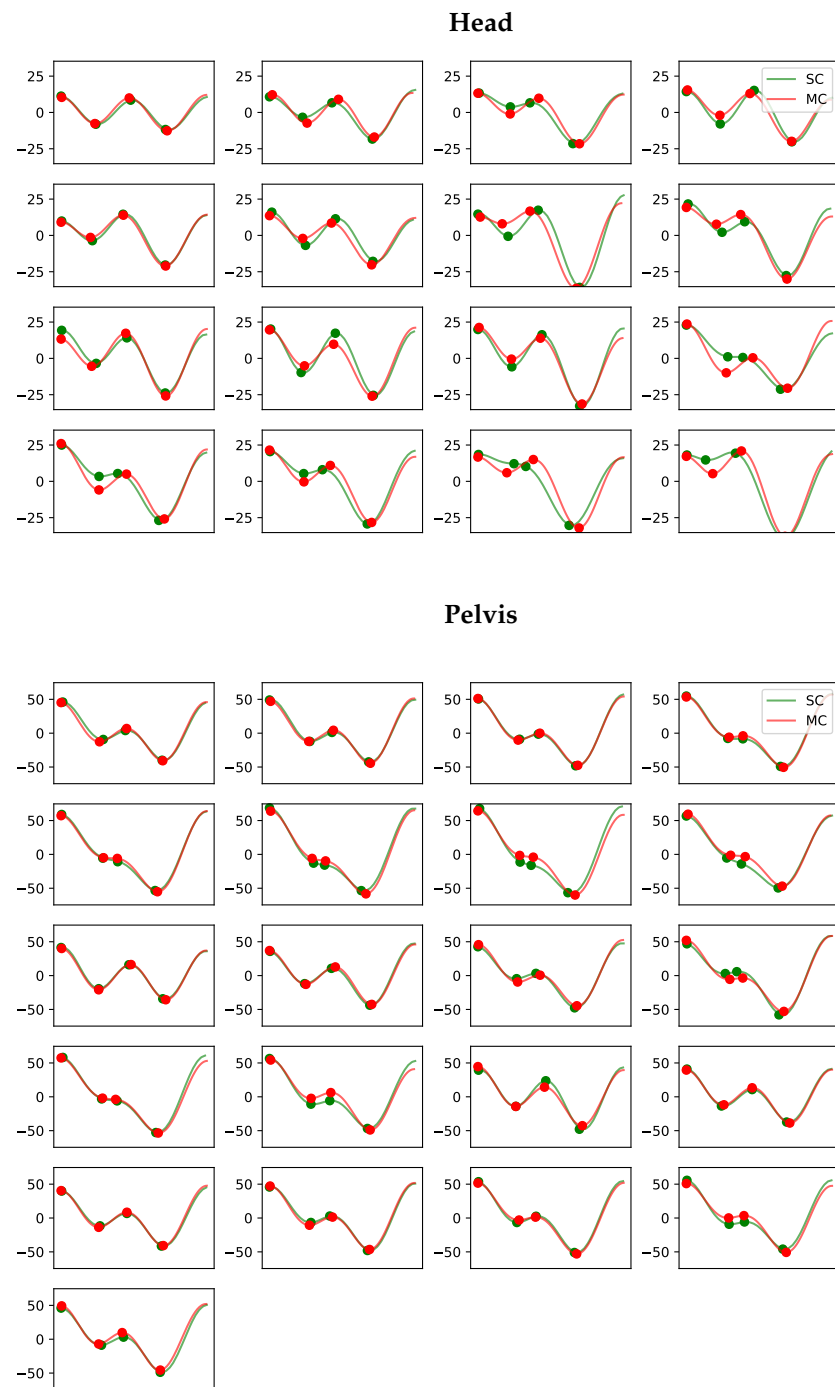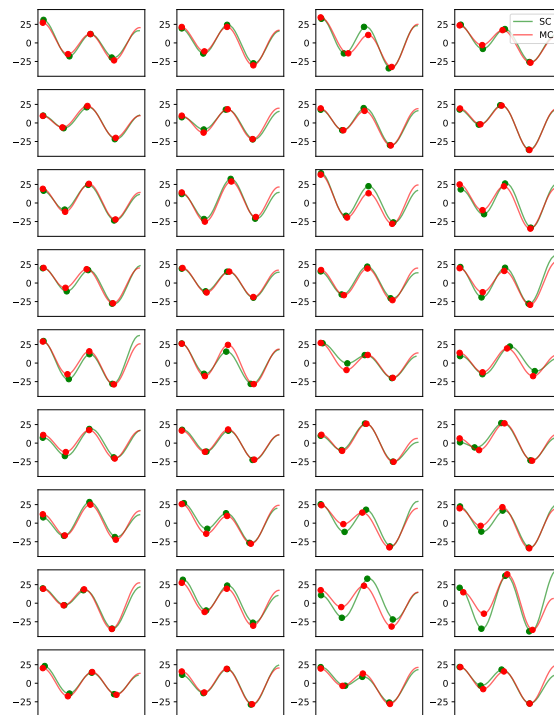
Head

Pelvis

Figure A19. Vertical displacement signals per stride for the head and pelvis for horse 20. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
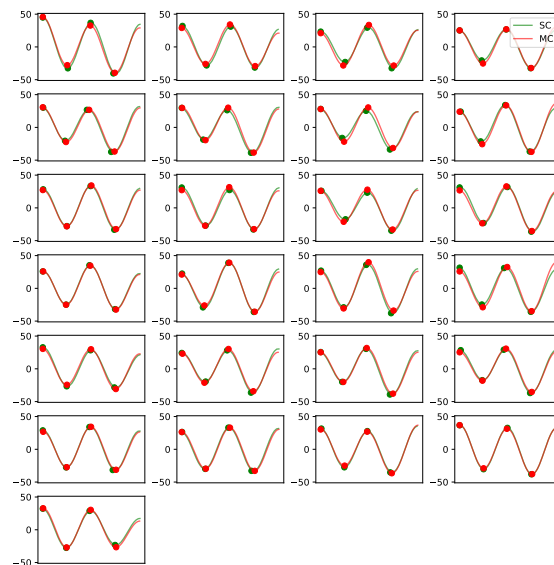
Head

Pelvis

Figure A20. Vertical displacement signals per stride for the head and pelvis for horse 21. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.

Head

Pelvis

Figure A21. Vertical displacement signals per stride for the head and pelvis for horse 22. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
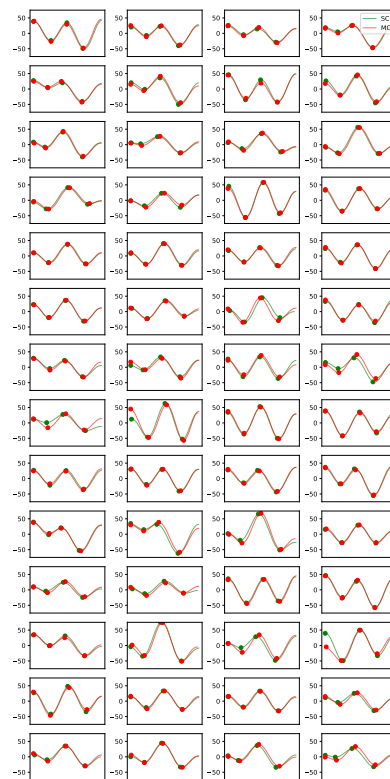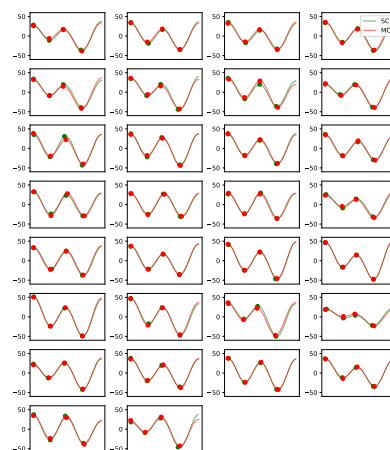
Head

Pelvis

Figure A22. Vertical displacement signals per stride for the head and pelvis for horse 23. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.
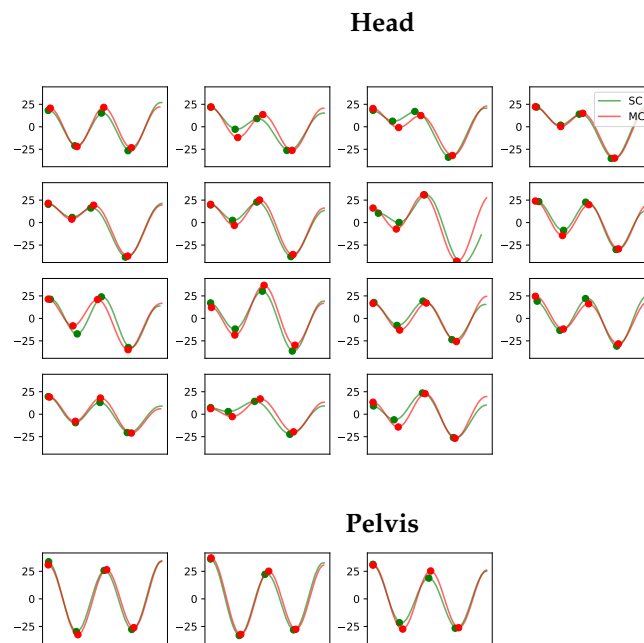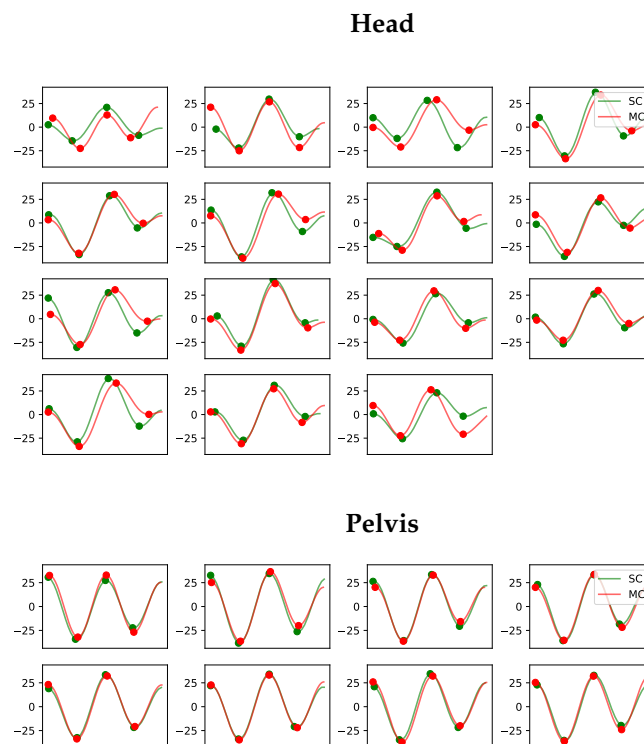
Head

Pelvis

Figure A23. Vertical displacement signals per stride for the head and pelvis for horse 24. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red. The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.

Head

Pelvis

Figure A24. Vertical displacement signals per stride for the head and pelvis for horse 25. Each subplot contains the matched stride segments for the single-camera markerless system (SC) in green and the multi-camera marker-based system (MC) in red). The y-axis shows the vertical displacement in millimeters. The four dots on each curve indicate the positions of the two peaks and two valleys extracted using the approach in Section 2.6.2.

References

1. Roepstorff, C.; Gmel, A.I.; Arpagaus, S.; Bragança, F.M.S.; Hernlund, E.; Roepstorff, L.; Rhodin, M.; Weishaupt, M.A. Modelling fore- and hindlimb peak vertical force differences in trotting horses using upper body kinematic asymmetry variables. *J. Biomech.* **2022**, *137*, 111097. [CrossRef] [PubMed]
2. Bell, R.; Reed, S.K.; Schoonover, M.; Whitfield, C.; Yonezawa, Y.; Maki, H.; Pai, F.P.; Keegan, K.G. Associations of force plate and body-mounted inertial sensor measurements for identification of hind limb lameness in horses. *Am. J. Vet. Res.* **2016**, *77*, 337–345. [CrossRef] [PubMed]
3. Buchner, H.H.; Savelberg, H.H.; Schamhardt, H.C.; Barneveld, A. Head and trunk movement adaptations in horses with experimentally induced fore- or hindlimb lameness. *Equine Vet. J.* **1996**, *28*, 71–76. [CrossRef] [PubMed]
4. Parkes, R.S.; Weller, R.; Groth, A.M.; May, S.; Pfau, T. Evidence of the development of 'domain-restricted' expertise in the recognition of asymmetric motion characteristics of hindlimb lameness in the horse. *Equine Vet. J.* **2009**, *41*, 112–117. [CrossRef]
5. Holcombe, A.O. Seeing slow and seeing fast: Two limits on perception. *Trends Cogn. Sci.* **2009**, *13*, 216–221. [CrossRef]
6. Keegan, K.G.; Dent, E.V.; Wilson, D.A.; Janicek, J.; Kramer, J.; Lacarrubba, A.; Walsh, D.M.; Cassells, M.W.; Schiltz, P.; Frees, K.E.; et al. Repeatability of subjective evaluation of lameness in horses. *Equine Vet. J.* **2010**, *42*, 92–97. [CrossRef]
7. Hammarberg, M.; Egenvall, A.; Pfau, T.; Rhodin, M. Rater agreement of visual lameness assessment in horses during lungeing. *Equine Vet. J.* **2016**, *48*, 78–82. [CrossRef]
8. Arkell, M.; Archer, R.M.; Guitian, F.J.; May, S.A. Evidence of bias affecting the interpretation of the results of local anaesthetic nerve blocks when assessing lameness in horses. *Vet. Rec.* **2006**, *159*, 346–349. [CrossRef]
9. Mccracken, M.J.; Kramer, J.; Keegan, K.G.; Lopes, M.; Wilson, D.A.; Reed, S.K.; Lacarrubba, A.; Rasch, M. Comparison of an inertial sensor system of lameness quantification with subjective lameness evaluation. *Equine Vet. J.* **2012**, *44*, 652–656. [CrossRef]
10. Rhodin, M.; Egenvall, A.; Andersen, P.H.; Pfau, T. Head and pelvic movement asymmetries at trot in riding horses in training and perceived as free from lameness by the owner. *PLoS ONE* **2017**, *12*, 1–16. [CrossRef]
11. Müller-Quirin, J.; Dittmann, M.T.; Roepstorff, C.; Arpagaus, S.; Latif, S.N.; Weishaupt, M.A. Riding Soundness—Comparison of Subjective With Objective Lameness Assessments of Owner-Sound Horses at Trot on a Treadmill. *J. Equine Vet. Sci.* **2020**, *95*, 103314. [CrossRef]
12. Pfau, T.; Noordwijk, K.; Caviedes, M.F.S.; Persson-Sjodin, E.; Barstow, A.; Forbes, B.; Rhodin, M. Head, withers and pelvic movement asymmetry and their relative timing in trot in racing Thoroughbreds in training. *Equine Vet. J.* **2017**, *50*, 117–124. [CrossRef] [PubMed]
13. Pfau, T.; Parkes, R.S.; Burden, E.R.; Bell, N.; Fairhurst, H.; Witte, T.H. Movement asymmetry in working polo horses. *Equine Vet. J.* **2016**, *48*, 517–522. [CrossRef] [PubMed]
14. Scheidegger, M.D.; Gerber, V.; Dolf, G.; Burger, D.; Flammer, S.A.; Ramseyer, A. Quantitative Gait Analysis Before and After a Cross-country Test in a Population of Elite Eventing Horses. *J. Equine Vet. Sci.* **2022**, *117*. [CrossRef] [PubMed]
15. Lopes, M.A.; Eleuterio, A.; Mira, M.C. Objective Detection and Quantification of Irregular Gait With a Portable Inertial Sensor-Based System in Horses During an Endurance Race—A Preliminary Assessment. *J. Equine Vet. Sci.* **2018**, *70*, 123–129. [CrossRef]
16. Kallerud, A.S.; Fjordbakk, C.T.; Hendrickson, E.H.S.; Persson-Sjodin, E.; Hammarberg, M.; Rhodin, M.; Hernlund, E. Objectively measured movement asymmetry in yearling Standardbred trotters. *Equine Vet. J.* **2020**, *53*, 590–599. [CrossRef]
17. Maliye, S.; Voute, L.C.; Marshall, J.F. Naturally-occurring forelimb lameness in the horse results in significant compensatory load redistribution during trotting. *Vet. J.* **2015**, *204*, 208–213. [CrossRef]
18. Pfau, T.; Witte, T.H.; Wilson, A.M. A method for deriving displacement data during cyclical movement using an inertial sensor. *J. Exp. Biol.* **2005**, *208*, 2503–2514. [CrossRef] [PubMed]
19. Keegan, K.G.; Yonezawa, Y.; Pai, P.F.; Wilson, D.A. Accelerometer-based system for the detection of lameness in horses. *Biomed. Sci. Instrum.* **2002**, *38*, 107–112.
20. Bosch, S.; Bragança, F.M.S.; Marin-Perianu, M.; Marin-Perianu, R.; van der Zwaag, B.; Voskamp, J.; Back, W.; van Weeren, R.; Havinga, P. EquiMoves: A Wireless Networked Inertial Measurement System for Objective Examination of Horse Gait. *Sensors* **2018**, *18*, 850. [CrossRef] [PubMed]
21. Hardeman, A.M.; Bragança, F.M.S.; Swagemakers, J.H.; van Weeren, P.R.; Roepstorff, L. Variation in gait parameters used for objective lameness assessment in sound horses at the trot on the straight line and the lunge. *Equine Vet. J.* **2019**, *51*, 831–839. [CrossRef]
22. Serra Bragança, F.M.; Rhodin, M.; van Weeren, P.R. On the brink of daily clinical application of objective gait analysis: What evidence do we have so far from studies using an induced lameness model? *Vet. J.* **2018**, *234*, 11–23. [CrossRef] [PubMed]
23. Aurand, A.M.; Dufour, J.S.; Marras, W.S. Accuracy map of an optical motion capture system with 42 or 21 cameras in a large measurement volume. *J. Biomech.* **2017**, *58*, 237–240. [CrossRef] [PubMed]
24. LeCun, Y.; Bengio, Y. Convolutional networks for images, speech, and time series. *Handb. Brain Theory Neural Netw.* **1995**, *3361*, 1995.
25. Fukushima, K. Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural Netw.* **1988**, *1*, 119–130. [CrossRef]
26. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]

27. Azhand, A.; Rabe, S.; Müller, S.; Sattler, I.; Heimann-Steinert, A. Algorithm based on one monocular video delivers highly valid and reliable gait parameters. *Sci. Rep.* **2021**, *11*, 1–10. [CrossRef]

28. Wang, J.; Sun, K.; Cheng, T.; Jiang, B.; Deng, C.; Zhao, Y.; Liu, D.; Mu, Y.; Tan, M.; Wang, X.; et al. Deep high-resolution representation learning for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 3349–3364. [CrossRef] [PubMed]

29. Li, C.; Ghorbani, N.; Broomé, S.; Rashid, M.; Black, M.J.; Hernlund, E.; Kjellström, H.; Zuffi, S. hSMAL: Detailed Horse Shape and Pose Reconstruction for Motion Pattern Recognition. *arXiv* **2021**, arXiv:2106.10102.

30. Hatherley, J.J. Limits of trust in medical AI. *J. Med. Ethics* **2020**, *46*, 478–481. [CrossRef]

31. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.

32. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.

33. Bhat, G.; Lawin, F.J.; Danelljan, M.; Robinson, A.; Felsberg, M.; Gool, L.V.; Timofte, R. Learning what to learn for video object segmentation. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 777–794.

34. Bragança, F.S.; Roepstorff, C.; Rhodin, M.; Pfau, T.; Van Weeren, P.; Roepstorff, L. Quantitative lameness assessment in the horse based on upper body movement symmetry: The effect of different filtering techniques on the quantification of motion symmetry. *Biomed. Signal Process. Control* **2020**, *57*, 101674. [CrossRef]

35. Keegan, K.G.; Yonezawa, Y.; Pai, F.; Wilson, D.A.; Kramer, J. Evaluation of a sensor-based system of motion analysis for detection and quantification of forelimb and hind limb lameness in horses. *Am. J. Vet. Res.* **2004**, *65*, 665–670. [CrossRef]

36. Pfau, T.; Boultbee, H.; Davis, H.; Walker, A.; Rhodin, M. Agreement between two inertial sensor gait analysis systems for lameness examinations in horses. *Equine Vet. Educ.* **2016**, *28*, 203–208. [CrossRef]

37. Bland, J.M.; Altman, D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *327*, 307–310. [CrossRef]

38. Keegan, K.G.; Kramer, J.; Yonezawa, Y.; Maki, H.; Pai, P.F.; Dent, E.V.; Kellerman, T.E.; Wilson, D.A.; Reed, S.K. Assessment of repeatability of a wireless, inertial sensor–based lameness evaluation system for horses. *Am. J. Vet. Res.* **2011**, *72*, 1156–1163. [CrossRef]

39. Wang, Y.; Li, J.; Zhang, Y.; Sinnott, R.O. Identifying lameness in horses through deep learning. In Proceedings of the ACM Symposium on Applied Computing, Virtual, 22–26 March 2021; pp. 976–985. [CrossRef]