

Properties of movement of 3D agents

Krzysztof Gorgolewski Maciej Komosinski Konrad Miazga
Krzysztof Rosiński Paweł Rychły

Technical Report RA-1/2019

Institute of Computing Science
Poznan University of Technology

Abstract

In this work we introduce eight well-defined properties of movement that can be calculated for morphologies of simulated agents, robots, or living creatures. The morphology of an agent is assumed to consist of connected parts. The properties are then calculated as aggregations of changes in position of these parts in time. We discuss the characteristics of these properties of movement, apply them as components of a dissimilarity measure, and demonstrate how they can be applied to the analysis of populations of simulated agents and their locomotion patterns. This report analyzes the dissimilarities between creatures with fixed morphologies and evolved controllers, and also between creatures with both morphologies and controllers evolved towards simple evolutionary goals. The correlation between the dissimilarity of creatures (measured as a weighted sum of the properties of the creatures) and the dissimilarity of their fitness is also discussed.

Contents

1	Introduction	1
2	Model of morphology	2
3	Numerical evaluation of locomotion patterns	3
3.1	Low-level raw data	3
3.2	High-level description of movement	3
3.3	Discussion of the properties	9
3.4	Illustrative examples	13
3.5	Dissimilarity measure and its properties	16
3.6	Correlations between the properties of movement	17
3.7	Estimating the importance of individual properties of movement	20
4	Properties of movement in specific evolutionary experiments	22
4.1	Estimating similarity between agents with common body morphology	22
4.2	Estimating similarity between agents with reference to the test environment	26
4.3	Estimating similarity between agents from many independent evolutionary runs	27
5	Summary and further work	34

1 Introduction

Over a million of organism species have been identified worldwide, each species having a specific body morphology and a more or less sophisticated nervous system. Many of these species have

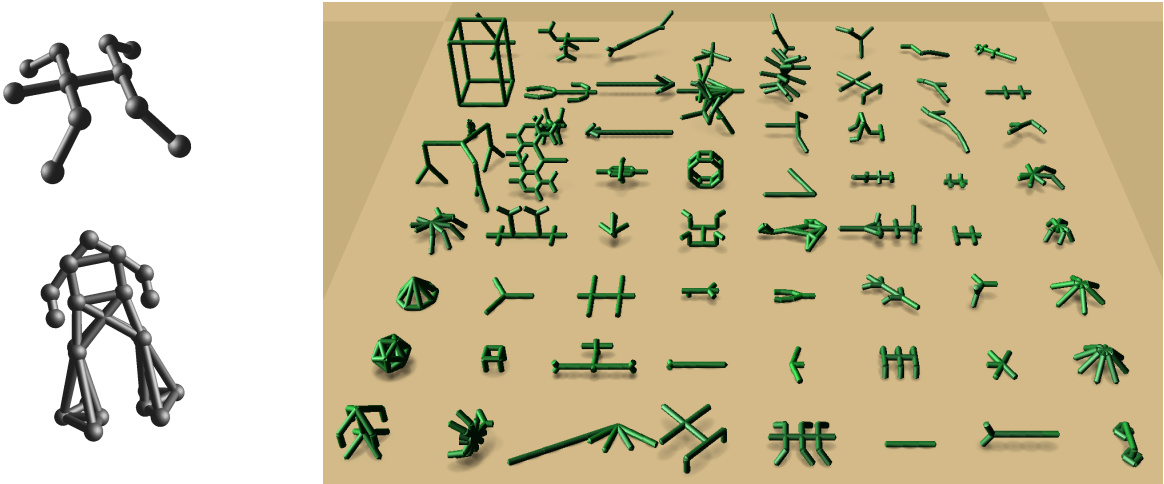


Figure 1: Sample 3D morphologies compatible with the model considered in this work. Left: two morphologies with visible parts (vertices) and their connections (edges). Right: an 8×8 sample of morphologies available in the Framsticks distribution [15].

once been a subject of a thorough observation and research. Scientists have gathered vast knowledge about existing organisms, thus it is possible to classify them with regard to various criteria.

However, imagine a new species is created with the use of genetic engineering, or is discovered, or a robot or a cyborg is designed and built. When one wants to classify such a creature, it turns out that the database of previously known creatures contains no information about it, and thus there is a need to gather or evaluate properties that describe this new organism.

When a person encounters an unknown creature, the first thing they do is perceiving it [21, 5]; in this process many qualities are investigated: the overall appearance, body morphology, and movement patterns [3, 23, 24]. While there are already some properties that are widely used for distinguishing between different body plans or types of movement (e.g., gait patterns), they are typically tailored towards specific kinds of creatures (e.g., quadrupeds, birds, or humans [16]) and they do not generalize well to all possible body plans and movement strategies. This work introduces a set of well-defined properties that describe movement patterns with scalar values, and demonstrates how these values can be used to compare and classify simulated creatures based on their locomotion and changes in body shape. Additional analyses of the properties introduced in this paper and related experiments are described in [11].

This work is a part of a larger undertaking started nearly 20 years ago to develop a set of similarity measures that concern various aspects of artificial and real organisms, in order to support human researchers in the analysis of large populations of such organisms. This is especially important because of the development of more and more sophisticated artificial intelligence agents, robots and artificial life forms. Large numbers of such entities cannot be investigated manually by humans, as it would require a lot of time and the results would be subjective and inconsistent. The aspects of similarity that have been earlier investigated include fitness, genetics [14], morphology [7, 8, 10, 6, 9], and neural network dynamics [12]; this work concerns behavioral traits.

2 Model of morphology

Movement properties introduced in this report can be used for animals, robots or simulated agents that consist of connected parts. The information about connections between parts is not used during the calculation of the properties. It is assumed that the fact that parts are connected influences the dynamics of their position and orientation in space, which is used

as a fundamental information when calculating the properties of movement. The supported morphology model is compatible with the legacy Framsticks [15, 14] ball-and-stick models [13], as illustrated in Fig. 1.

3 Numerical evaluation of locomotion patterns

3.1 Low-level raw data

Each simulation of a creature provided three series of values. Each series consisted of the following values, respectively:

- The current central point (x, y, z) – can be the center of mass, the geometric center, or the center of the bounding box of the creature.
- The dispersion in the xy plane: d_{xy} .
- The dispersion in the z dimension: d_z .

Each central point is described by three coordinates corresponding to the three dimensions. Calculating dispersions requires some elementary data processing. Since creatures are made of distinguishable elements called parts (connected by joints), we define creature “dispersion” in the appropriate plane as the weighed standard deviation [17] of its body parts. Let c be the central point of a creature and P be the set of all body parts. The dispersion in the xy plane is evaluated as follows (dispersion in the z dimension is calculated analogously):

$$d_{xy} = \sqrt{\frac{\sum_{p \in P} \text{weight}(p) * ((p.x - c.x)^2 + (p.y - c.y)^2)}{\sum_{p \in P} \text{weight}(p)}} \quad (1)$$

where $\text{weight}(p)$ is the importance of part p . In the following experiments, $\text{weight}(p)$ is the mass¹ of part p . This reflects the idea that the movement of the heavy body parts, such as a head, affects the values of the properties more than the movement of the light body parts, such as a finger. Alternatively, if one were interested in a purely visual or geometrical description of body movements, the value of $\text{weight}(p)$ could be set to 1 for all parts p .

3.2 High-level description of movement

These raw data series are hard to interpret directly; they form a basis for higher-level descriptors of movement. We have introduced eight properties describing different qualities of locomotion:

1. Average error of linear correlation of position in the xy plane (*err_line_xy*). Low value indicates movement similar to a straight line.
2. Horizontal oscillation factor (*var_dis_xy*) – standard deviation of dispersion in xy divided by mean dispersion in xy .
3. Vertical oscillation factor (*var_dis_z*) – standard deviation of dispersion in z divided by mean dispersion in z .
4. Vertical-to-horizontal oscillation ratio (*sd_dis_z_xy*) – mean ratio of dispersion in z to dispersion in xy .
5. Mean speed in xyz (*inst_speed*).
6. Spectral Flatness Measure (*sfm*) defined as a geometrical mean of frequency domain of xyz speed divided by its arithmetical mean.

¹In the simulations reported here, the mass of each part was always equal to the number of coincident joints.

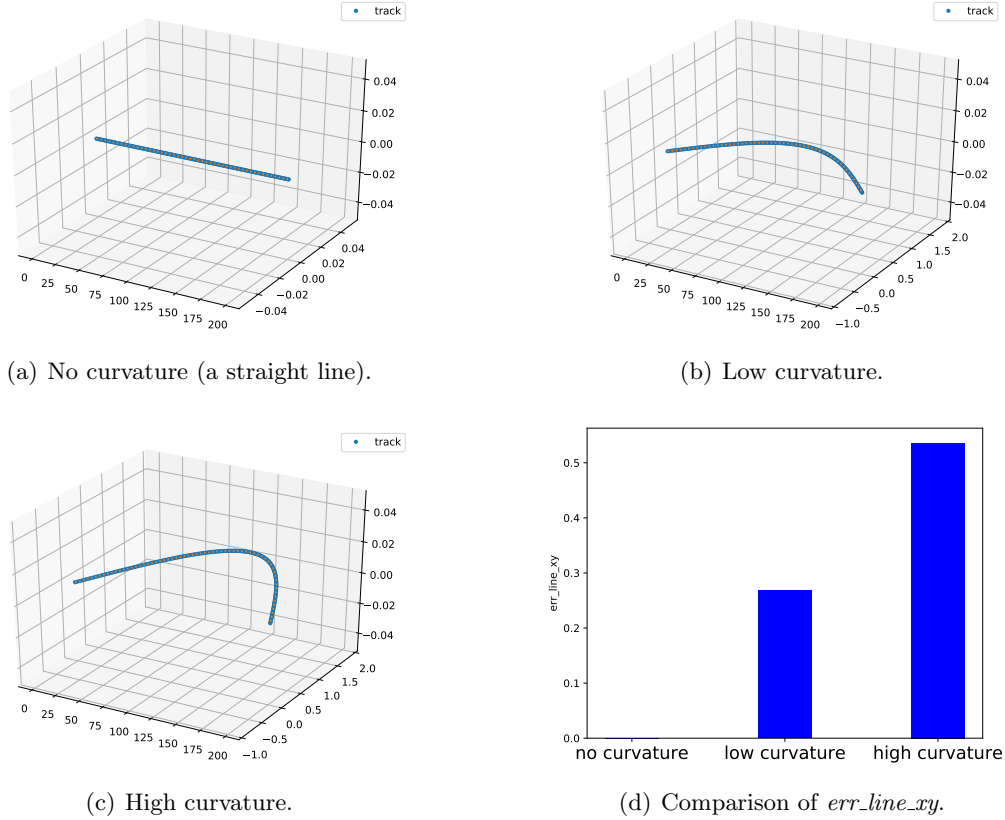


Figure 2: The relationship between the path of movement and the value of err_line_xy .

7. The frequency of the highest amplitude of the xyz speed (f_max).
8. The maximal correlation of the xyz speed signal with its offset ($autocorr_max$).

The definition and calculation of these properties refer to basic concepts in linear algebra [20, 4, 19], statistics and regression analysis [1, 18] and signal analysis [22, 2]. Short descriptions of the properties and the corresponding equations are presented below.

In the following equations, D_{xy} and D_z are dispersion series in planes xy and z respectively, and σ denotes standard deviation. Average values are overlined, for example $\overline{d_{xy}}$ means the average dispersion in the xy plane. Let us define $E(a, b)$ as the Euclidean distance between points a and b , c_i being the i -th sample of a creature body center series C . Sets X, Y, Z contain series of the center points in a particular dimension.

1. The average error of two-dimensional euclidean regression in the xy plane (err_line_xy) tells us whether a creature changes its direction during simulation or oscillates around the path of movement.

To calculate this property it is necessary to find an euclidean regression line $y = ax + b$ [27]. That is the line for which a sum of the squares of distances from points of trajectory is minimal. A distance ε_i between a point of trajectory (x_i, y_i) and the regression line may be calculated in the following way:

$$\varepsilon_i = \left| \frac{y_i - (ax_i + b)}{\sqrt{a^2 + 1}} \right| \quad (2)$$

The average error is calculated as a mean value of these distances.

To illustrate this property three trajectories with different shape were generated. Fig. 2 illustrates a comparison of these trajectories.

2. The horizontal oscillation factor evaluates creature movement dynamics in the xy plane (var_dis_xy). A creature that has a large horizontal oscillation factor tends to put a lot of effort in the horizontal limb moves like walking, crawling and snake-style swimming. The horizontal oscillation factor may be evaluated as follows:

$$var_dis_xy = \frac{\sigma(D_{xy})}{\bar{d}_{xy}} \quad (3)$$

3. The vertical oscillation factor (var_dis_z) may also be considered a simple indicator of a type of movement. It is very similar to the horizontal oscillation factor, the difference is that instead of the xy plane, the z dimension is used. Intuitively, a large vertical oscillation factor may indicate a jumping type of movement, while smaller values may characterize crawling creatures. The vertical oscillation factor may be calculated in the following way:

$$var_dis_z = \frac{\sigma(D_z)}{\bar{d}_z} \quad (4)$$

4. The vertical-to-horizontal oscillation ratio ($sd_dis_z_xy$) may be interpreted as a combination of the vertical oscillation factor and the horizontal oscillation factors. Each of the previously presented factors may be interpreted separately. In order to obtain a relative combination of these two properties we divide one by the other. The denominators used in var_dis_z and var_dis_xy properties have been dropped. The equation describing the vertical-to-horizontal oscillation ratio is presented below:

$$sd_dis_z_xy = \frac{\sigma(D_z)}{\sigma(D_{xy})} \quad (5)$$

To illustrate these three properties, two straight trajectories with different dispersions series were generated. Fig. 3 illustrates relations between dispersion of movement and values of the properties. Dispersion of movement in the first example changed more dynamically in the xy plane, Fig. 3(a). Dispersion of the second movement changed more dynamically in the z dimension, Fig. 3(b).

5. The mean instantaneous speed ($inst_speed$) in xyz may be evaluated as follows:

$$\bar{v} = \frac{\sum_{i=2}^{|C|} E(c_i, c_{i-1})}{|C|} \quad (6)$$

Before the control systems of creatures were activated, there was a wait-for-stabilization period so that creature speed in the first time step t_0 was always 0.

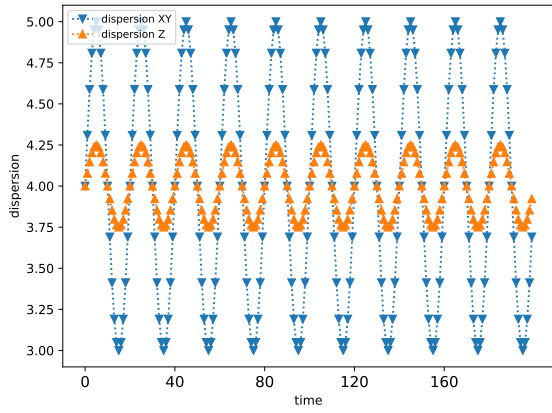
To illustrate mean speed property, a straight trajectory with a sinusoidal speed was generated as shown in Fig. 4(b). In this example speed oscillates around the calculated mean value ($inst_speed = 0.49765$).

6. Spectral Flatness Measure (sfm) is one of two properties that uses the discrete Fourier transform. This property technically evaluates the dynamics of a period of movement. If the speed of a creature may be presented as a sinusoid then it is expected to have a low sfm value. Creatures having irregular speed will have higher sfm value. The equation used for evaluating the Spectral Flatness, along with predecesing operations, is presented below. The first step is to calculate the displacement vector:

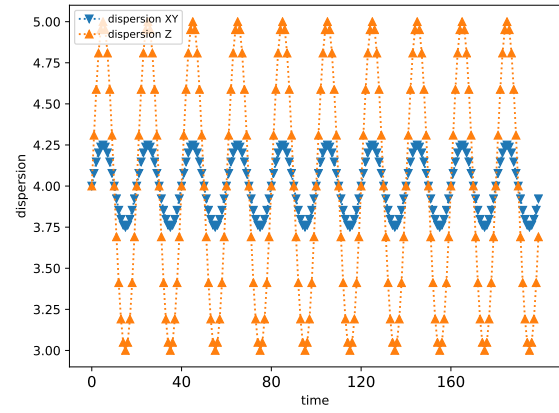
$$\Delta C = \{c'_i : c'_i = E(c_{i+1}, c_i); c_i \in C, i = 1..|C| - 1\} \quad (7)$$

Secondly, the Fourier transformation (DFT) is performed on the vector ΔC

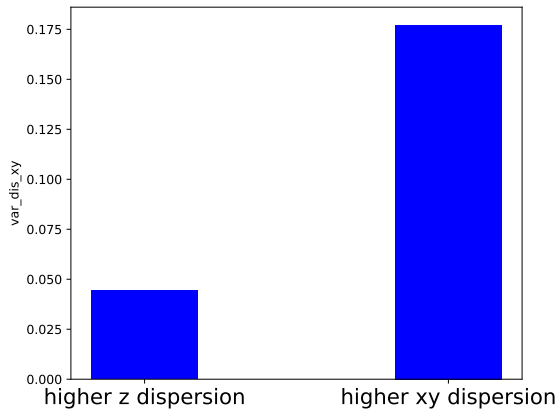
$$F = \text{DFT}(\Delta C) \quad (8)$$



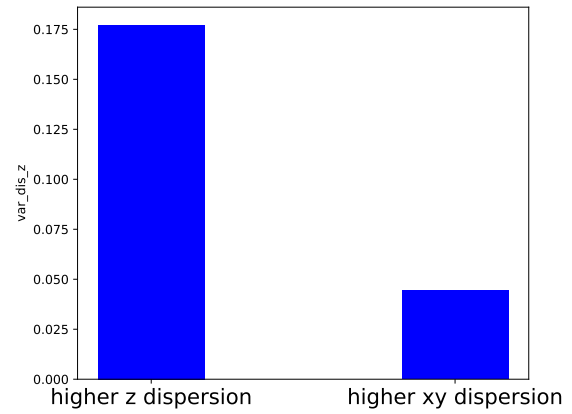
(a) Higher dispersion in the xy plane.



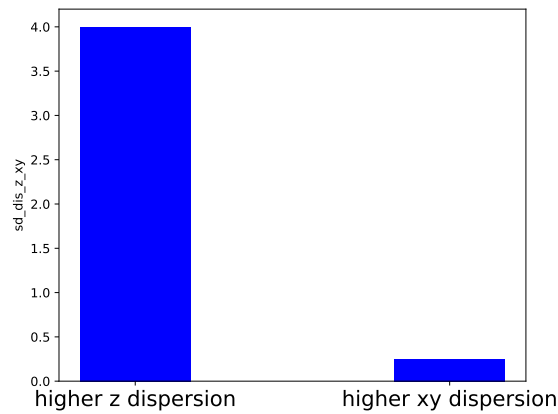
(b) Higher dispersion in the z dimension.



(c) Comparison of var_dis_xy .

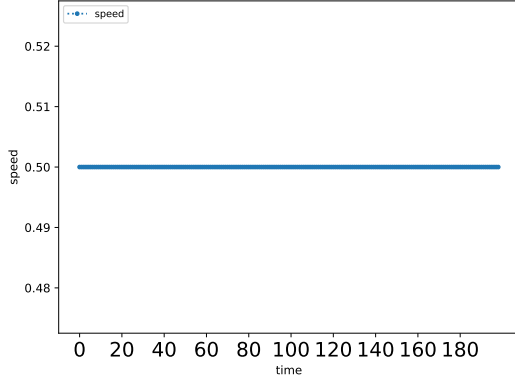


(d) Comparison of var_dis_z .

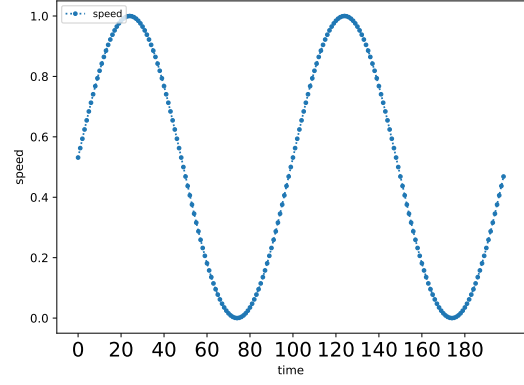


(e) Comparison of $sd_dis_z_xy$.

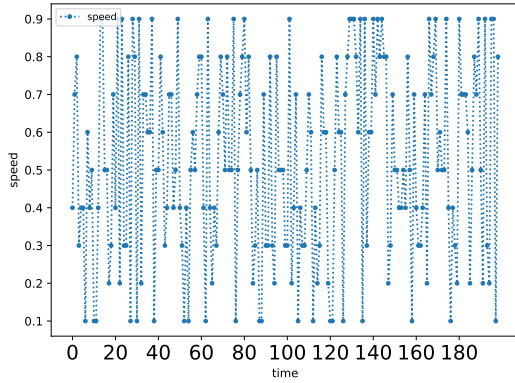
Figure 3: The relationship between dispersion values in the xy plane and the z dimension and the values of var_dis_xy , var_dis_z and $sd_dis_z_xy$.



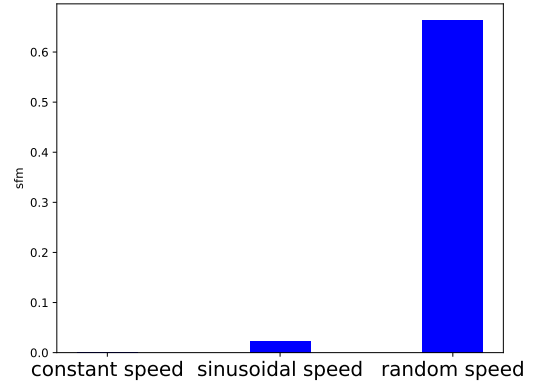
(a) Constant speed.



(b) Sinusoidal speed.



(c) Random speed.



(d) Comparison of *sfm*.

Figure 4: The relationship between the shape of speed in time and the value of *sfm*.

Finally, the *sfm* is calculated.

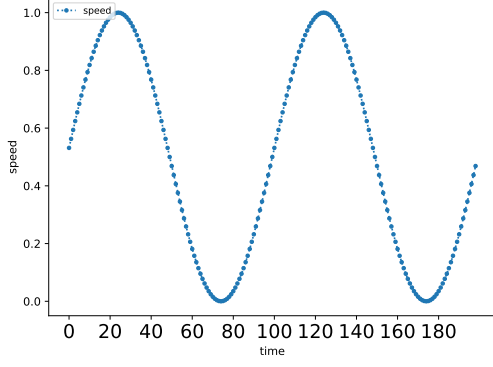
$$sfm = \frac{e^{\overline{\ln(F)}}}{\overline{F}} \quad (9)$$

Fig. 4 illustrates the relation between the shape of the speed function and value of spectral flatness measure. If the function of instantaneous speed is constant in time as in Fig. 4(a), or sinusoidal in time as in Fig. 4(b), the value of *sfm* is lower. If the speed of an agent is irregular as in Fig. 4(c), then the value of the *sfm* is higher.

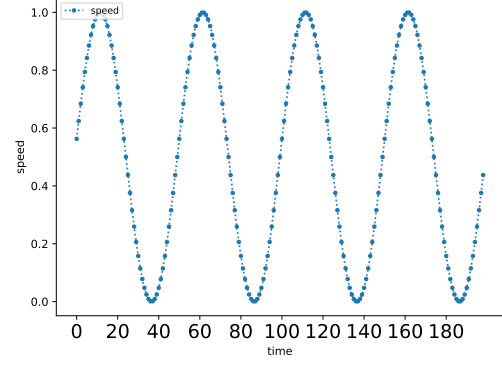
7. The most significant frequency (the frequency with the highest amplitude, f_{max}) is a simple property that provides information about the most prominent frequency in the spectrum of the creature's function of speed in time:

$$f_{max} = \arg \max(F) \quad (10)$$

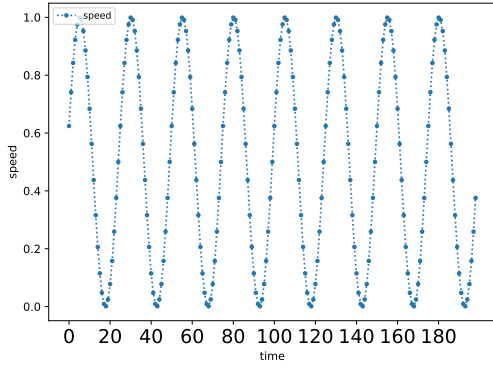
The $\arg \max$ function returns the index of the highest value in a vector. This value represents the most significant frequency (other than 0 Hz, technical aspects are not presented in order to simplify the description). The f_{max} property is illustrated in Fig. 5.



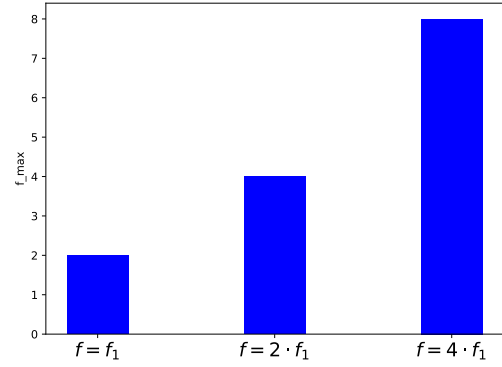
(a) Sinusoidal speed with frequency f_1 .



(b) Sinusoidal speed with frequency $2 \cdot f_1$.



(c) Sinusoidal speed with frequency $4 \cdot f_1$.



(d) Comparison of f_{max} .

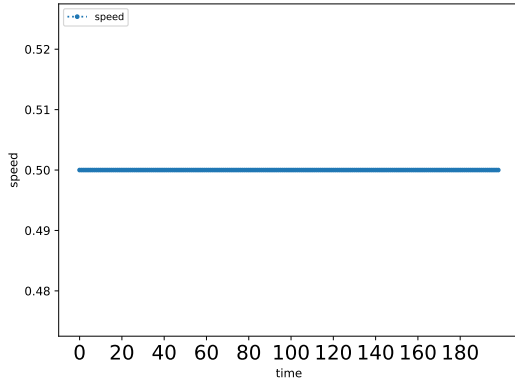
Figure 5: The relationship between the frequency of the sinusoidal signal of creature's speed and the value of f_{max} .

8. The maximal correlation of the xyz speed signal with its offset ($autocorr_{max}$) can be calculated according to the algorithm:

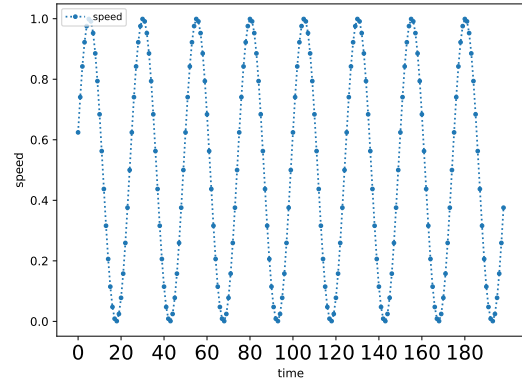
```

1:  $Speed \leftarrow \text{vector}(|C| - 1)$ 
2: for  $i = 0..|C| - 3$  do
3:    $Speed[i] = E(C[i], C[i + 1])$ 
4: end for
5:  $AutoCorr \leftarrow \text{vector}(|C|/2 - 1)$ 
6:  $acLen \leftarrow |AutoCorr|$ 
7:  $AutoCorr[0] \leftarrow 1$ 
8: for  $i = 1..acLen - 1$  do
9:    $AutoCorr[i] \leftarrow$ 
      $\text{corrcoef}(Speed[0 : acLen], Speed[i : acLen + i])$ 
10: end for
11:  $firstBetter \leftarrow 0$ 
12: for  $i = 1..acLen - 1$  do
13:   if  $AutoCorr[i] > AutoCorr[i - 1]$  then
14:      $firstBetter \leftarrow i$ 
15:     break
16:   end if
17: end for
18:  $autocorr_{max} \leftarrow$ 
      $\text{max}(AutoCorr[firstBetter : acLen])$ 

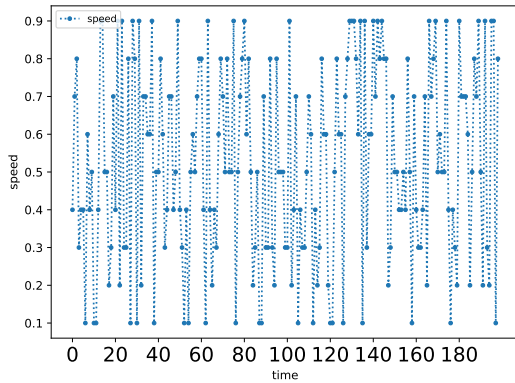
```

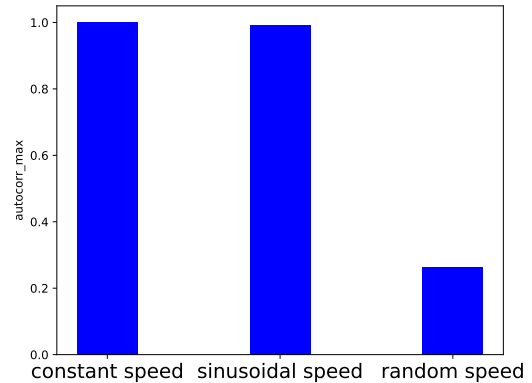
(a) Constant speed.



(b) Sinusoidal speed.



(c) Random speed.



(d) Comparison of *autocorr_max*.

Figure 6: The relationship between agent’s speed and the value of *autocorr_max*.

In this pseudocode, C is the vector containing the center point of the creature for each time frame. Vectors containing elements from a specific range are represented by the base vector name and a corresponding range in square brackets separated by a colon, for example $Speed[x : y]$ refers to a subvector containing elements from vector $Speed$ with indices from x to y (excluding y). The `corrcoef` function in line 9 calculates the correlation coefficient between the original and the shifted vector. The `for` loop in lines 11-17 determines *firstBetter* – the index of the first local minimum in the *AutoCorr* vector, plus one. The offsets lower than this index are then excluded from the search for the highest correlation (line 18).

The *autocorr_max* property is illustrated in Fig. 6.

The values of all the proposed properties have been visualized in Fig. 7 for agents from the Scorpion (land), Fastlizard (land) and Star (land) populations. These populations were evolved (with fixed morphology) starting from the creatures with the corresponding names; the genotypes of Scorpion and Fast lizard are available in the Framsticks distribution [15], and Star was a three-pointed regular star body defined by the $XX(XX,XX)$ genotype.

3.3 Discussion of the properties

It would be desirable for all the properties introduced in the previous section to have the quality of *scalability*. By scalability we mean the invariance of the values of properties in face of the increased resolution of sampling of the points on the body of a creature. If the data series for the creature contained samples from two evenly spaced points for each body part instead of

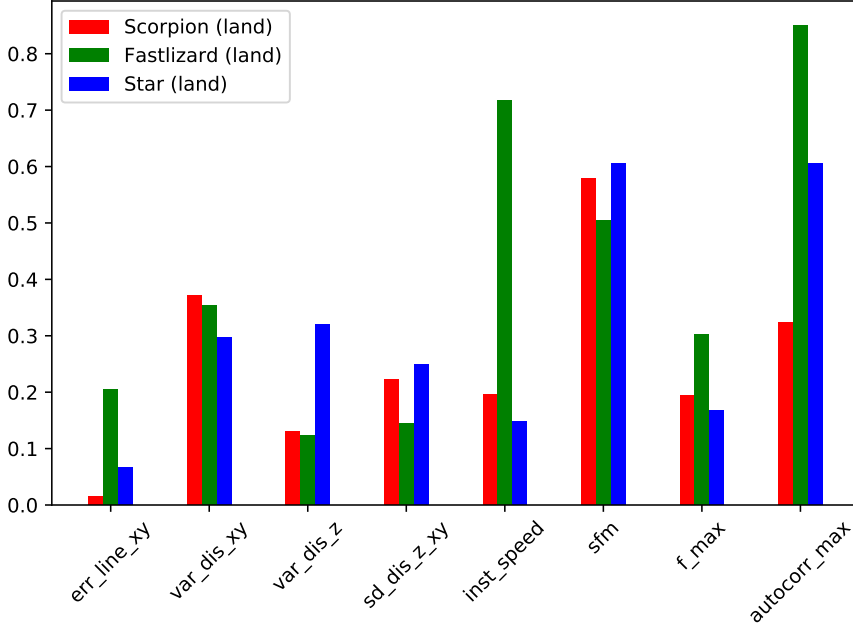


Figure 7: Values of the proposed properties for three different populations (i.e., “species”) of creatures.

just one point, the values of the properties should not change, or at least change only slightly, with the magnitude of changes decreasing to zero as the resolution grows to infinity. This quality would facilitate comparing the values of the properties between creatures of free-form blob-like morphologies that are not necessarily made of easily distinguishable parts (modules). Although such analysis was not performed in this report, it should be done in the future, and the properties that are not scalable should be modified to achieve this quality.

Below, we include an additional discussion of the properties introduced in the previous section.

err_line_xy As shown earlier, *err_line_xy* can distinguish between the trajectories of different curvature. It is not however indifferent to the other parameters of the movement, such as the speed of movement or the length of the considered signal. If the trajectory of the movement is linearly approximable and the oscillations around the linear path are not growing in time, the value of *err_line_xy* should asymptotically converge towards some specific value. An example of such a behavior of the property for a sinusoidal trajectory with different speeds of movement can be seen in Fig. 8. However, if the path of a creature cannot be approximated with a straight line, the value of *err_line_xy* may not converge towards any value as the speed of the creature (or the length of the signal) grows.

Another potential problem with this property comes from the assumption that the movement of the creature takes place mostly in the horizontal plane. While that assumption is likely to be true for walking creatures, it leads to ignoring an entire dimension of movement for the flying and swimming creatures. This problem could be solved by extending the property to a three-dimensional linear approximation of the path of movement *err_line_xyz*. That change, however, would mean that for the walking creatures the shape of the terrain could influence the exact value of *err_line_xy*, e.g. the value would be higher for a hilly terrain than it would be for a flat surface.

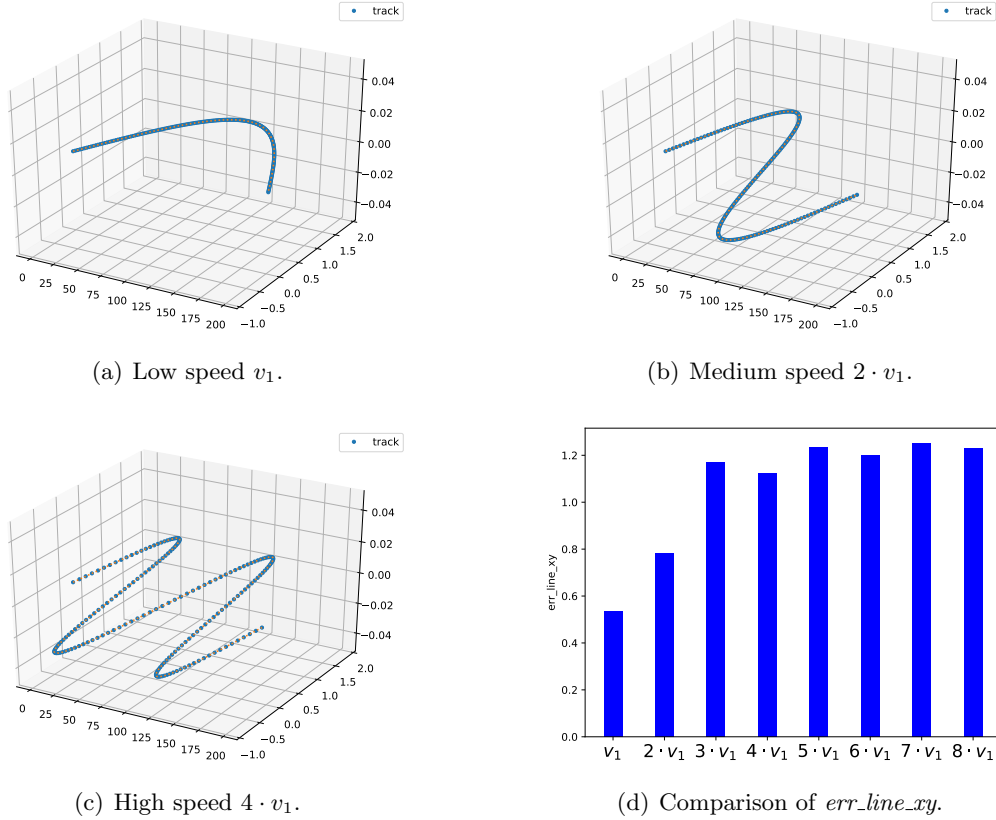


Figure 8: The relationship between the speed of movement and err_line_xy for a fixed sinusoidal path of movement. Analogous results can be obtained for a constant speed of movement and proportionally longer signals.

var_dis_xy, var_dis_z, sd_dis_z_xy The dispersion of a creature used for calculating these properties weighs each body part according to the physical weight of that part. That means that two similar creatures with different weight distributions will be assigned different dispersion values – the calculated dispersion value for a creature with a heavy core and light limbs will be lower than it would be if the core of that creature was light and its limbs were heavy. Therefore, var_dis_xy and var_dis_z assess not only the amount of movement in a given plane/dimension, but the effort required to perform that movement.

The downside of this definition of dispersion is that two creatures with – seemingly – the same movement may have different values of the dispersion-based properties. It is not clear which approach would be more useful in distinguishing between different classes of movement – perhaps both the weighted and the unweighted versions these properties should be considered when comparing two types of movement.

inst_speed This property calculates the average speed between two consecutive measurements of the position of the creature, and therefore it can be affected by the frequency with which the considered signal is being sampled.

It is important to note that high instantaneous speed does not imply high overall speed measured as displacement divided by time, although the opposite is true – in fact the instantaneous speed determines the upper bound on the possible speed value as measured over longer periods of time. Two examples of high instantaneous speed yielding low overall speed are the circular motion (the creature regularly returns to its starting position) and the Brownian motion (the creature randomly changes the direction of the movement, and so its displacement from the initial position changes very slowly and irregularly). It could therefore be worth to extend the

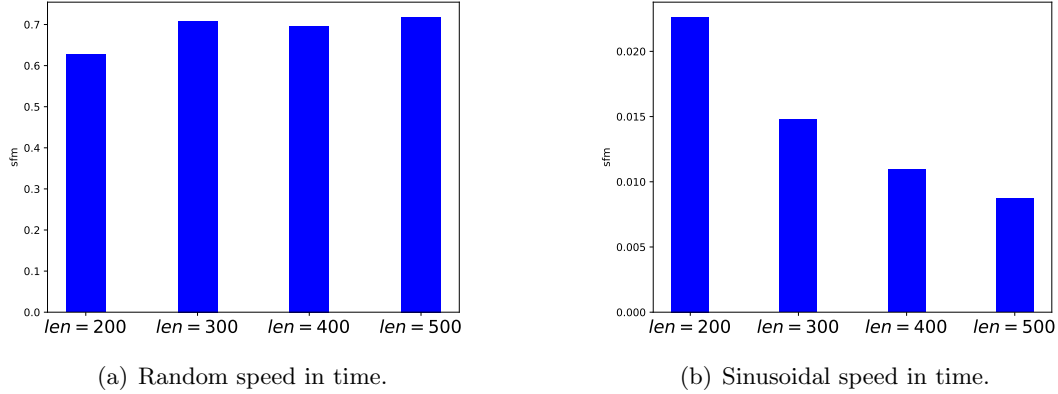


Figure 9: The relationship between the length of the signal and sfm for the random and the sinusoidal speed in time. While the random speed in time has a similar amount of power in all spectral bands, the sinusoidal speed in time contains only one frequency.

inst_speed property to the average speed of the creature as a function of the length of the time window. The shape of such a function would give one a better understanding of the dynamics of the creature: for a constant linear motion such function would be constant, for the circular motion the function would be clearly periodic, and for the Brownian motion the value of the function would be irregularly decreasing. Although such a function would not be limited to a single value and as such it would not be very concise, a number of new properties based on its shape could be introduced.

sfm The exact value of sfm may depend on the length of the signal, especially for the more regular movement with lower sfm values. The longer the considered signal, the wider the frequency spectrum, and as the spectrum for a more regular movement will contain only a limited number of spikes, increasing its length will decrease both the arithmetic and the geometric means. The geometric mean will however decrease faster than the arithmetic mean, and so as the length of the signal increases, the value of sfm will tend to decrease. This effect can be seen in Fig. 9.

f_max In order to find the frequency with the highest amplitude, a Fourier transform must be used to generate the spectrum of the signal. However, before applying the Fourier transform to the signal, one of a number of different window functions should be applied to that signal in order to improve the properties of the resulting spectrum. Two of the most important properties that can be modified by applying a window function to a signal are the resolution and the dynamic range of the signal's spectrum. The high resolution means that the strength of the part of the spectrum corresponding to some frequency will not leak to the nearby frequencies – it will however leak to the frequencies that are farther away. On the other hand, the high dynamic range means that the spectrum of a frequency present in the signal will have its leakage to the farther frequencies heavily reduced, at the cost of leaking to the nearby frequencies.

This trade-off, known as the spectral leakage, forces us to choose between the high resolution and the high dynamic range of the spectrum. The high dynamic range is important mostly for distinguishing weaker frequencies present in the signal. For the purpose of the *f_max* property we are however interested in the strongest frequency, and so we are willing to sacrifice the dynamic range of the spectrum for its high resolution. Therefore, for the purpose of our experiments we have used the rectangular window when preparing the signal for the Fourier transform. The use of that window guarantees us the highest possible resolution of the resulting spectrum.

For some of the considered creatures the highest component of the signal of the speed in

time is the constant speed. In the case of these creatures, the non-zero frequencies are often very low and show poor signal to noise ratio. To avoid assigning undeservingly high values of f_max to these creatures, the creatures for which the highest amplitude of the non-zero frequencies is lower than 0.002 are assigned $f_max = 0$.

The frequency spectrum of the speed signal can be multimodal, and multiple maxima can have similar amplitudes. In such cases, the value of f_max may depend on tiny differences in simulation (such as different initial conditions or noise) and thus may change radically between evaluations. This is an undesired effect of using the maximum function. A better way of calculating the dissimilarity between the frequency spectrums of the speed signals of two creatures could involve the EMD (earth mover’s distance) measure [26]. Such a measure would however be unfeasible as a descriptor of movement of each individual creature, as it describes the relation between two creatures instead of the property of one of them.

autocorr_max In order to calculate the value of *autocorr_max*, a vector of the autocorrelation values for various shifts must first be prepared. There are two main ways to calculate this vector.

In the first approach the entire available signal is always used to calculate the autocorrelation for each shift. That means that when calculating the autocorrelation for some shift $s \in [1; L/2]$, we compare the segments of the original signal taken from ranges $[0; L - s]$ and $[s; L]$, where L is the length of the full signal. The advantage of this approach is that each value of the autocorrelation vector is calculated based on the full available data. The disadvantage is that the bigger the shift, the lower the precision of the calculated value (i.e., the correlation is being calculated for shorter segments) is going to be.

On the other hand, in the second approach the lengths of the segments of the signal that are being compared (and therefore the precision of the resulting autocorrelation value) are fixed. In this approach, when calculating the autocorrelation for some shift $s \in [1; L/2]$, we compare the segments of the signal taken from ranges $[0; L/2]$ and $[s; L/2 + s]$. Although this allows for a fixed precision of the calculations for the entire vector, that precision (and so the amount of data used per shift) will on average be lower than in the first approach.

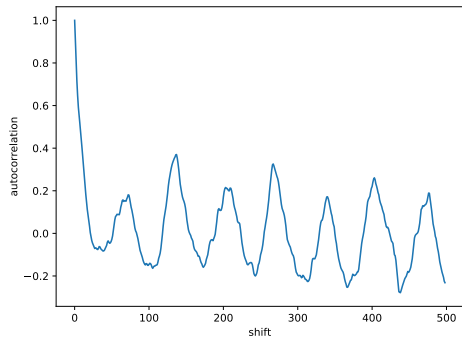
The comparison of the full autocorrelation vectors calculated using both approaches can be seen in Fig. 10. Although there are some differences between the results obtained with the two approaches described above, no clear repeatable characteristic patterns can be observed for any of the approaches. It is therefore difficult to unambiguously determine which of those two methods is better. For the purpose of this report, the second approach, which assumes the fixed lengths of the compared vectors, is used.

3.4 Illustrative examples

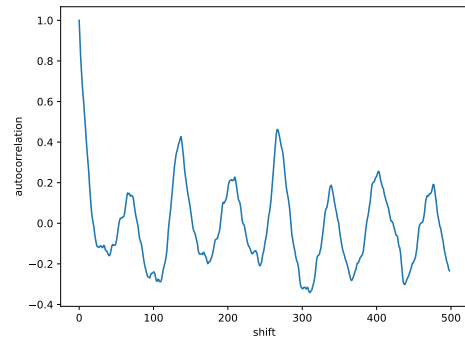
Fig. 11 presents a comparison of some selected walking and swimming creatures available in the Framsticks distribution [15].

The first of the selected walking creatures, the SlowSmellingPuller, performs a very systematic dragging motion – a single front leg repeatedly drags the rest of the body forward. Although this creature does not move very fast (it has a medium value of *inst_speed*), it moves in a straight line (as it has a symmetrical body which is being dragged forward by a centered front leg) and so its value of the *err_line_xy* is very close to zero. The high regularity of its movement leads to a high value of *autocorr_max*. Although the period of this creature’s movement cycle is comparable to the period of the movement cycle of other creatures such as Frog#2, one cycle of the movement of SlowSmellingPuller consists of two phases – the first one being the pulling motion, and the second being the rebound which happens when the front leg is being lifted. For that reason the value of f_max for SlowSmellingPuller is about twice as high as it is for Frog#2.

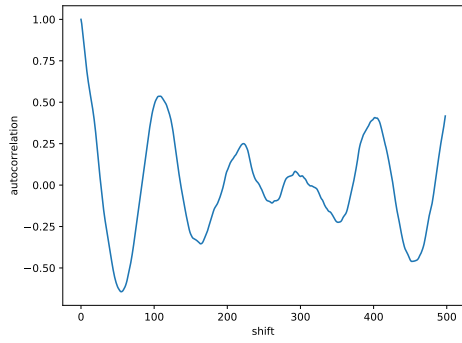
In contrast to the SlowSmellingPuller, the Speedy creature has a very high value of *err_line_xy*. The reason for this is the curved path of the movement of the creature, which combined with



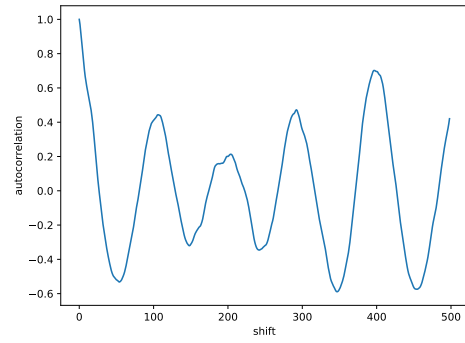
(a) First approach – creature #1.



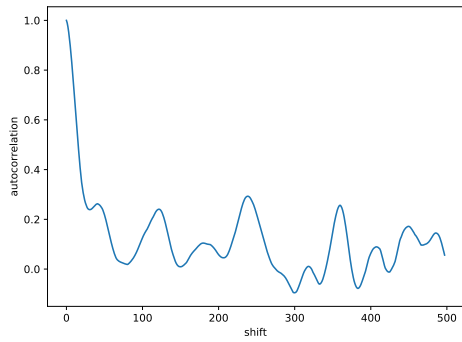
(b) Second approach – creature #1.



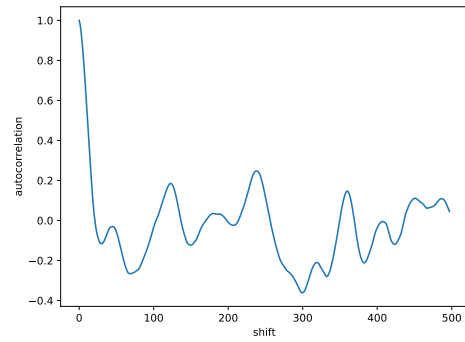
(c) First approach – creature #2.



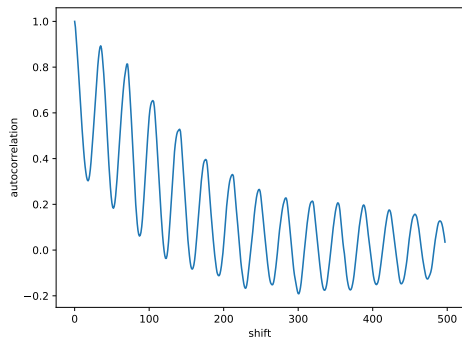
(d) Second approach – creature #2.



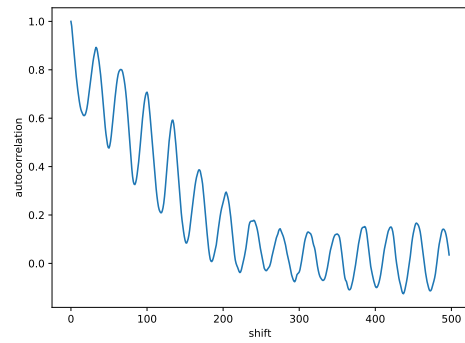
(e) First approach – creature #3.



(f) Second approach – creature #3.

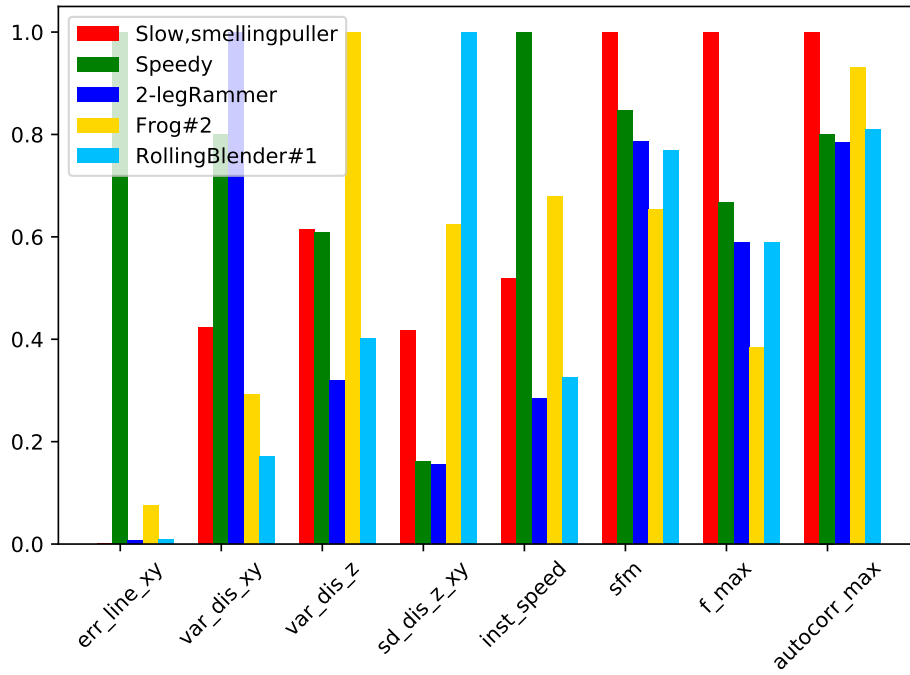


(g) First approach – creature #4.

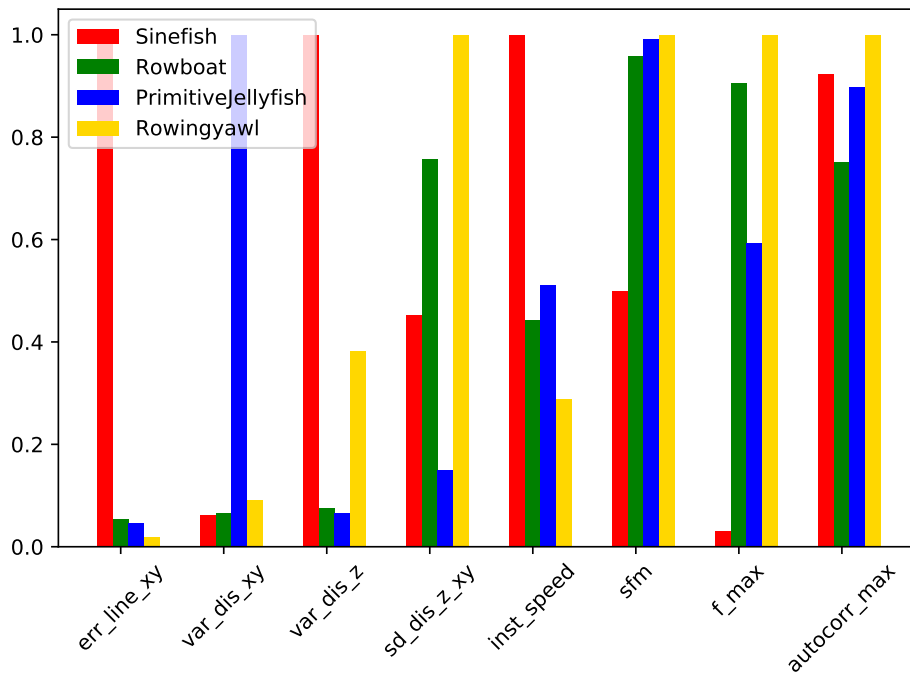


(h) Second approach – creature #4.

Figure 10: The autocorrelation vectors for five different creatures, calculated with two approaches.



(a) A comparison of the selected walking creatures.



(b) A comparison of the selected swimming creatures.

Figure 11: Comparisons of some selected walking (*walking.gen*) and swimming (*swimming.gen*) creatures available in the Framsticks distribution [15]. The values have been normalized according to the maximum values of the properties among the creatures presented on a given plot.

the high value of *inst_speed* leads to the high average error of the linear approximation of that path.

The other three creatures: 2-legRammer, Frog#2 and RollingBlender#1, obtain very high values of – respectively – *var_dis_xy*, *var_dis_z* and *sd_dis_z_xy*. This can also be explained by the specifics of their movement. 2-legRammer is composed of two “legs” connected with a horizontal stick, which causes the creature to bounce left and right with each consecutive step. Frog#2, while flat, performs a leaping motion which causes it to throw its front higher upwards then its back while jumping, leading to a perceived elongation in the vertical dimension. The body of RollingBlender#1 consists of four prongs, two on each side of the body, with the first two attached orthogonally to the other two. The creature performs a rolling motion, which in combination with its shape leads to a comparable dispersion in both xy-plane and z-dimension, therefore leading to a high value of *sd_dis_z_xy*.

Similar observations can be made for the swimming creatures. Sinefish is fast but does not (at least in the initial stages of its simulation) follow a linear path, so both its *inst_speed* and *err_line_xy* are high. PrimitiveJellyfish shows strong outwards-inwards movement (like an umbrella, opening and closing) which is reflected in the high value of *var_dis_xy*. Rowingyawl balances its movement in xy-plane and z-dimension through a repeated rowing motion, therefore it obtains a high value of *sd_dis_z_xy*. The most intriguing value which can be observed in Fig. 11(b) is however the high value of *var_dis_z* for Sinefish. Such a high value of that property is intriguing, as the Sinefish creature is flat and does not show a lot of vertical movement. The little vertical movement that it shows does, however, vary a lot in time when compared to the average value of the dispersion of the Sinefish in the vertical dimension. This is confirmed by an average value of *sd_dis_z_xy* for that creature – even though Sinefish shows a high variation of its vertical dispersion, that dispersion itself must still be relatively small when compared to the horizontal dispersion of this creature.

3.5 Dissimilarity measure and its properties

The dissimilarity measure between agents *a* and *b* may be calculated as presented in (11). The set of properties is defined as *P*, and thus *a_i* and *b_i* refer to the numerical evaluation of the *i*-th property of agent *a* and *b*, respectively. Some properties may provide more information about agent similarity, on the other hand some may provide none, therefore there is a need to adjust the impact of each property on the final dissimilarity value. In the presented equation, *w_i* defines the weight of the *i*-th property.

$$d(a, b) = \sum_{i \in P} w_i |a_i - b_i| \quad (11)$$

Each agents properties of movement may be evaluated without prior knowledge about other agents in the pool. Such quantitative information is objective since it does not depend on the context. Therefore it is easy to verify that the dissimilarity measure (11) is a metric. Assuming that *A* is the set of agents, a measure may be considered a metric if and only if the following statements are true for any *i, j, k* ∈ *A*:

- (a) $d(i, j) \geq 0$
- (b) $d(i, j) = d(j, i)$
- (c) $d(i, j) = 0 \Leftrightarrow i = j$
- (d) $d(i, k) \leq d(i, j) + d(j, k)$

Condition (a) is always fulfilled because of the absolute value in (11). The dissimilarity value could become negative if some (or all) weights would be negative, yet since it is obvious that these weights should always be positive (or at least non-negative), this will not happen.

Statement (b) is true because the absolute value of a subtraction of the two values is the same regardless of which value is the minuend and which is the subtrahend.

Condition (c) states that the dissimilarity value may be 0 if and only if the agent is compared to an identical agent. Although (c) may not always be true because agents with different morphologies and/or control systems may yield the same values of movement properties, we can make this condition true by adding the following two restrictions:

$$a = b \Leftrightarrow \forall_{i \in P} a_i = b_i \quad (12)$$

$$\forall_{i \in P} w_i > 0 \quad (13)$$

In (12) we assume that only the properties from the P set are used to differentiate agents (“closed world assumption”), so agents with identical movement properties are considered indistinguishable. To ensure that the agents with distinguishable values of the properties from P remain distinguishable, the restriction (13) forbids zero weights. Moreover, since in (13) we assume all weights are positive and, in (11), the absolute value of the difference between the values of the properties is used, a dissimilarity value of 0 cannot be achieved by compensating the deviation in value of one property with the deviation in value of another property. Therefore, under these two restrictions, condition (c) is true.

The last condition (d) is fulfilled because the agent dissimilarity measure is a sum of one-dimensional Euclidean distances and the Euclidean distance itself is a metric: if (d) is fulfilled for every element of the sum, it will also be true for the total sum.

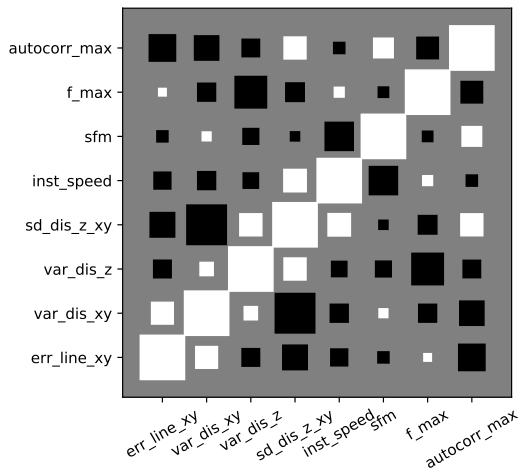
In conclusion, the proposed dissimilarity measure is a metric on the properties of movement of the creatures. Under the additional assumption (12), we can also consider it to be a metric on the creatures themselves.

3.6 Correlations between the properties of movement

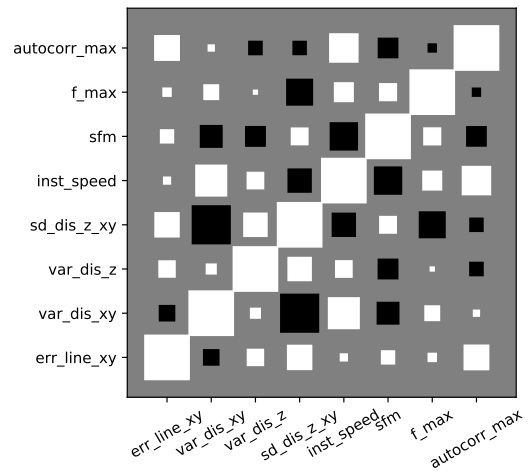
The properties were proposed after performing a large set of observations on diversified types of creatures. The output values describe creature movement patterns in a quantitative manner. Yet it was hard to forecast what are the relations between themselves. A good method to verify it is to test the correlation for each pair of properties. If some relations are observed, additional actions may be taken, for instance a high value of correlation between two properties may indicate that one of them is redundant and therefore can be ignored as it does not provide any additional information.

Experiments were performed separately for three different species of creatures: *fastlizard*, *scorpion* and *star*, with each species evolved in two different environments: land and water; six unique populations were considered in total. This way, it was possible to investigate whether the correlations between the values of properties observed in the population depend on the specific creatures, or are they inherent to the properties themselves – although the size of the experiment does not allow for any meaningful statistical reasoning, some of the more obvious trends may be clear upon visual examination. For every population, each agent had its movement patterns evaluated in a quantitative manner according to the equations presented earlier in this section. The results were then arranged into eight vectors – each corresponding to one property of movement. In the next step, the Pearson correlation coefficient was calculated for each pair of vectors. The results are presented in the form of matrices in Fig. 12.

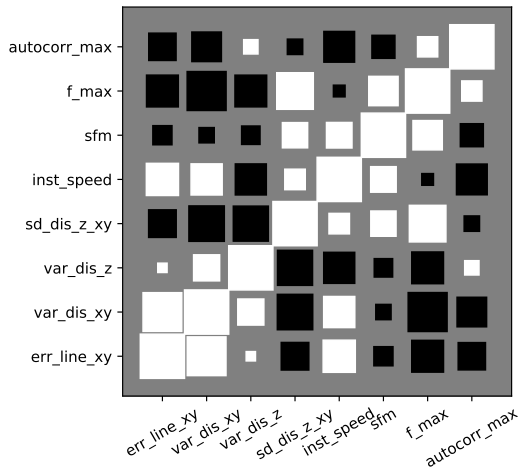
As expected, the highest values appear on the diagonal – they correspond to the correlation between properties and themselves. For most of the pairs of properties, the absolute correlation strengths are reasonably high, which suggests that the properties are not fully independent, as they often synergize with each other (positive correlation) or require trade-offs (negative correlation). However, in most cases these interdependencies do not appear to be inherent to the properties, but rather an effect of applying specific morphologies (species) to different environments. This observation is based on the fact that there are not many patterns that repeat for matrices in Fig. 12.



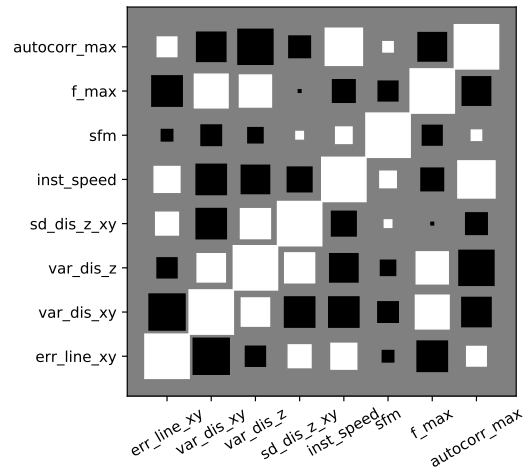
(a) Fastlizard (land).



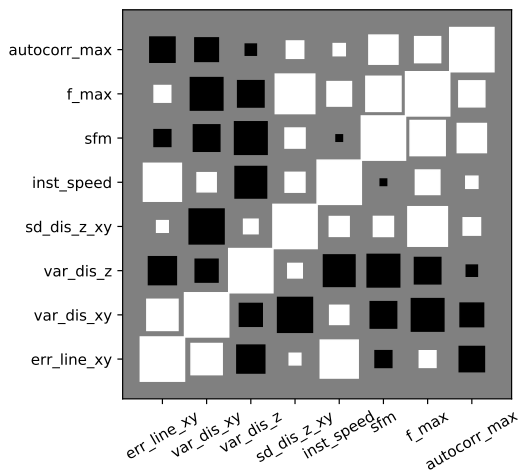
(b) Fastlizard (water).



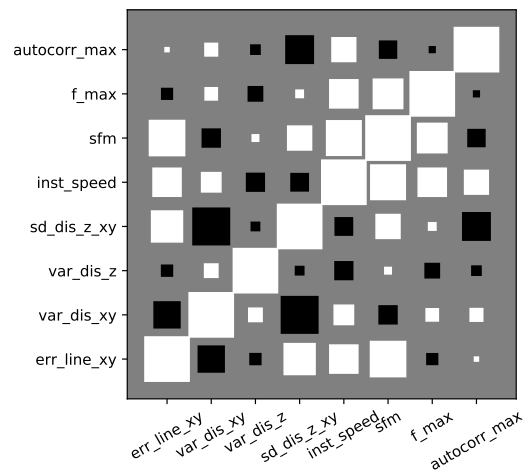
(c) Scorpion (land).



(d) Scorpion (water).

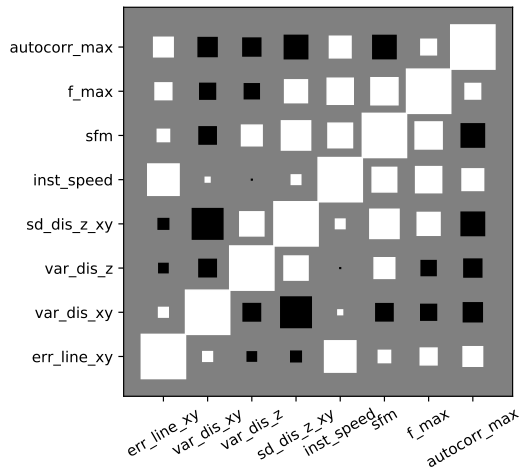


(e) Star (land).

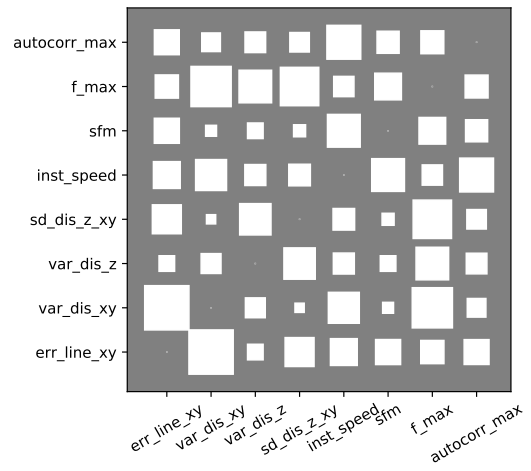


(f) Star (water).

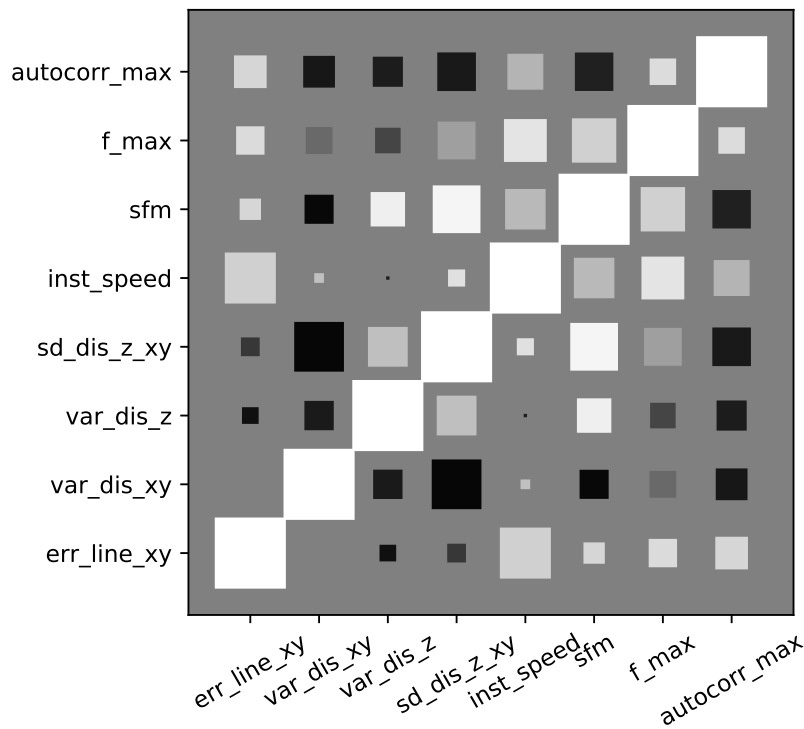
Figure 12: Movement property correlation matrix for different populations.



(a) Movement property correlation matrix for the combination of six populations.



(b) Variance of the movement property correlations between six populations.



(c) Movement property correlation matrix for the combination of six populations. This plot merges information from 13(a) and 13(b): transparency of the boxes shown in (a) reflects the interpopulation variance of the correlations shown in (b).

Figure 13: A summary of the property correlations for six populations (three different species: *fastlizard*, *scorpion*, and *star*, both on land and in water).

Fig. 13(a) presents the correlation matrix for the combined six aforementioned populations, while Fig. 13(b) presents the matrix of variation between the values of the correlations in these six populations. Low variation of the correlation values between different populations for some pairs of the movement properties may suggest that these properties may be inherently interdependent, in a way which does not depend on a specific morphology of the creatures (it may however still depend on the general evolutionary task towards which the creatures were evolved). The movement property correlation matrix with the box transparency scaled by the interpopulation variance of the correlations can be seen in Fig. 13(c).

The lowest variations occur for the pairs: $var_dis_xy - sd_dis_z_xy$, $sd_dis_z_xy - sfm$, and $var_dis_xy - sfm$. The first pair $var_dis_xy - sd_dis_z_xy$ shows relatively strong negative correlation, which can be explained by the formulas (3) and (5) with which these two properties are calculated: $\sigma(D_{xy})$ is a numerator of var_dis_xy , while simultaneously being a denominator of $sd_dis_z_xy$. The second pair $sd_dis_z_xy - sfm$ shows a positive correlation, which suggests that creatures which dynamics are dominated by the movement in the xy-plane (lower $sd_dis_z_xy$) tend to perform more regular movement (low sfm) than the creatures which show more movement in the vertical direction (e.g. by jumping). The last pair $var_dis_xy - sfm$ shows relatively weak negative correlation, low variance of which could be attributed to the low variance of the two aforementioned pairs of properties.

For some pairs of the properties, the variance is especially high. An example of such a pair of properties is $err_line_xy - var_dis_xy$ which appears to always correlate positively for the land populations, and negatively for the water populations. Although the number of the examined populations is too low to draw meaningful conclusion, the fact that the value of the $err_line_xy - var_dis_xy$ correlation is always positive for the land populations and negative for the water populations may suggest that the value of this correlation could be used to distinguish between the populations evolved in different environments.

3.7 Estimating the importance of individual properties of movement

Calculating distance based on the raw values of the properties may not yield the expected results as the scales on which these properties operate will not be the same across the whole set of the properties. It is therefore advised for the values of each property to be standardized. However, after the standardization the relative importance of each property will be equal when calculating the dissimilarity of two creatures. This can be seen as advantageous, although in practice, depending on the context, non-equal weights could also be seen as beneficial. An example of such a situation would be a population where value of one of the properties stays relatively constant – in this scenario the standardization of that property’s values could lead to an unfair magnification of small, meaningless differences between creatures. In a perfect scenario, the standardization of values of each property should be based on the set of all possible creatures, which unfortunately is impossible. Below, one of the possible ways to assign weights to each property is presented.

In order to calculate the weights of the properties based on their “differentiating power”, we determine the coefficient of variation over the set of average property values for different agent populations. Let us define C as the set of populations (classes), while P is the set of properties. The average property value matrix M_{avg} may be calculated in the following way:

$$\forall_{i=1..|P|} \forall_{j=1..|C|} M_{avg}[i][j] = \bar{c}_i : c_i \in C[j] \quad (14)$$

After the matrix is calculated, the coefficient of variation is determined for each column to estimate the diversity of the values of each property across all examined populations. The values of the coefficient of variation are then normalized (which causes the least important property to be assigned zero weight) and then rescaled to sum up to one. The visualization of the matrix and the coefficients of variation can be seen in Fig. 14. A high diversity value (top

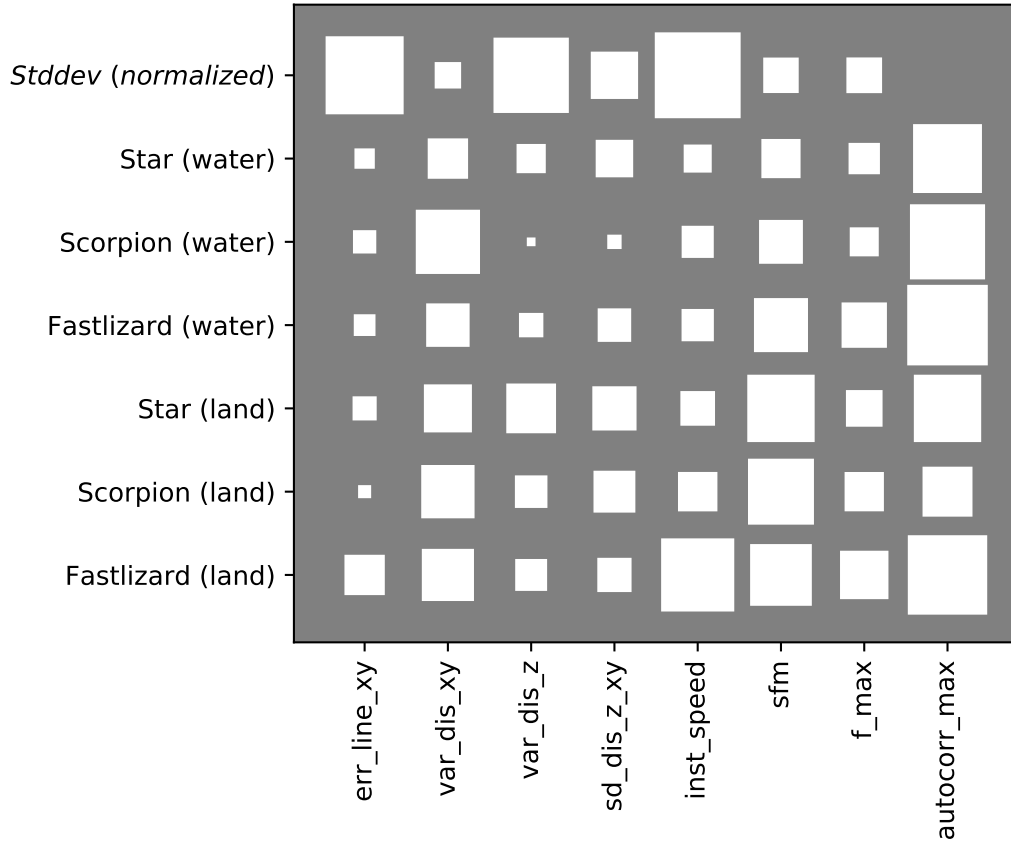


Figure 14: Matrix of the average values of properties in each population. The first row represents normalized values of the coefficient of variation for the properties (i.e., how strongly do the values vary relative to their mean).

Property	Pop. std. dev.	Norm. pop. std. dev.	Weight
<i>err_line_xy</i>	0.813	0.820	0.254
<i>var_dis_xy</i>	0.302	0.040	0.012
<i>var_dis_z</i>	0.775	0.763	0.237
<i>sd_dis_z_xy</i>	0.465	0.289	0.089
<i>inst_speed</i>	0.930	1.000	0.310
<i>sfm</i>	0.377	0.155	0.048
<i>f_max</i>	0.379	0.158	0.049
<i>autocorr_max</i>	0.276	0.000	0.000

Table 1: Movement property diversity and proposed weights.

row) indicates that the related property (column) should receive a high weight. Low diversity value indicate that the values of the property tends to be similar for different populations.

Table 1 presents the weights that were calculated for the set of creatures taken from six populations: Fastlizard, Scorpion, and Star, each of those in both water and land variants. The weights are clearly diversified, with the mean speed in *xyz* having the biggest impact on the dissimilarity measure with the weight of 31%; considering that there are 8 properties this is a very high value. Other important properties are *err_line_xy* (25.4%) and *var_dis_z* (23.7%). The maximum value of the autocorrelation *autocorr_max* has received a weight of 0%, which effectively excludes it from the set of the considered properties.

The proposed method of selecting weights of the properties is not optimal in the sense of maximization of the cohesion and separation of the clusters of creatures from different populations. The only effect the weighting of the properties has in the space of the original dimensionality is stretching of the dimensions. Weighting of the properties can however affect the low-dimensional projections of the full set of creatures by changing the relative distances between creatures. Therefore, the proposed method allows one to amplify the component of the variability present in the data – the component that distinguishes different populations. This method preserves the variability present within each population, but to a lesser degree than it would be the case if the weights of all properties were kept equal.

4 Properties of movement in specific evolutionary experiments

Going a step farther from the comparing the average values of the properties for different populations, a visualization of the distribution of creatures of different types in the space of the proposed properties could allow for evaluating the validity of the proposed set of properties. Should the proposed properties allow for distinguishing between different populations, this would suggest that the proposed set of the properties is sufficient. Unfortunately, due to the high number of the properties, the populations of creatures cannot be visualized directly in their original dimensionality.

To allow the visual comparison between different populations, we use the multidimensional scaling (MDS) to reduce the number of dimensions to three and two. In order to perform MDS, the matrix of distances between all creatures is first calculated. Then, the algorithm calculates the new coordinates for all the creatures in the new, low-dimensional space, while simultaneously preserving the original distances between them to the highest possible degree.

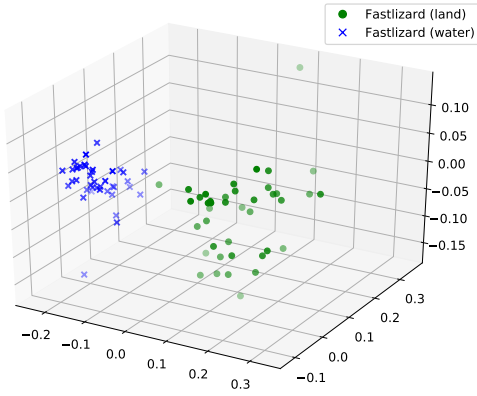
In the experiments below we compare the results obtained for two different weight profiles. In the first approach the weights all of properties are set to be equal, while in the second approach these weights are calculated specifically for the considered set of creatures following the procedure described in Sect. 3.7. The mean silhouette values [25] are also computed (in the original space of all properties) in order to quantitatively evaluate whether the proposed method of setting weights is beneficial. The mean silhouette describes the degree to which the observations (i.e., creatures) are located near the center of mass of their respective clusters (i.e., populations) and, simultaneously, away from other clusters. The silhouette measure calculated for a single creature takes the values in the range $[-1, 1]$, where -1 indicates that the creature is more similar to creatures from other populations than to its own respective population, while the value of 1 indicates that the creature is a perfect representative of its own population. Therefore, the mean silhouette value reflects how easy it is to distinguish the populations from each other.

Additional experiments where the movement properties introduced in this work are applied for the analysis of populations of evolved agents are described in [11].

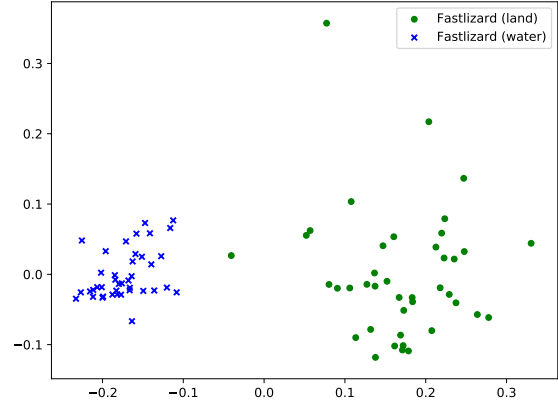
4.1 Estimating similarity between agents with common body morphology

In this experiment agents were aggregated into test populations so that each test population contained only agents with a similar body morphology. There were three test populations, each corresponding to one type of body morphology:

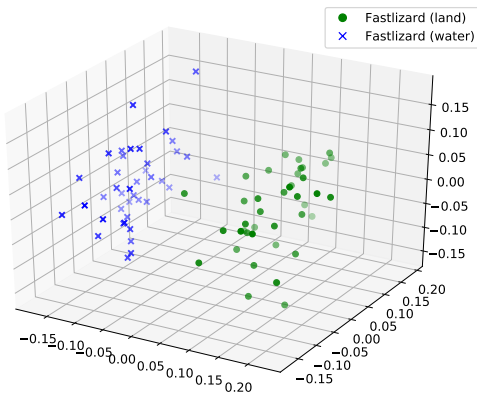
- (a) Fastlizard population – aggregates Fastlizard (land) and Fastlizard (water).
- (b) Scorpion population – aggregates Scorpion (land) and Scorpion (water).
- (c) Star population – aggregates Star (land) and Star (water).



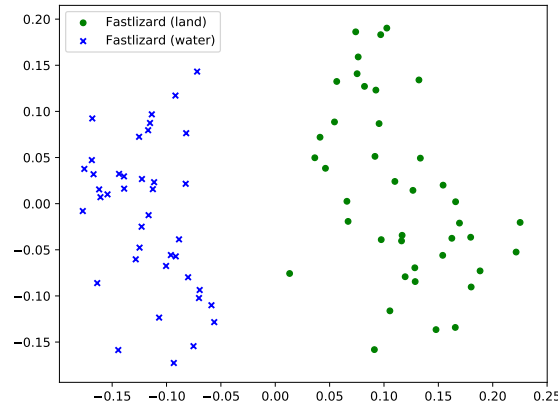
(a) Specific weights, 3D space (68.7% of the total variance preserved).



(b) Specific weights, 2D space (63.8% of the total variance preserved).



(c) Equal weights, 3D space (53.1% of the total variance preserved).



(d) Equal weights, 2D space (44.7% of the total variance preserved).

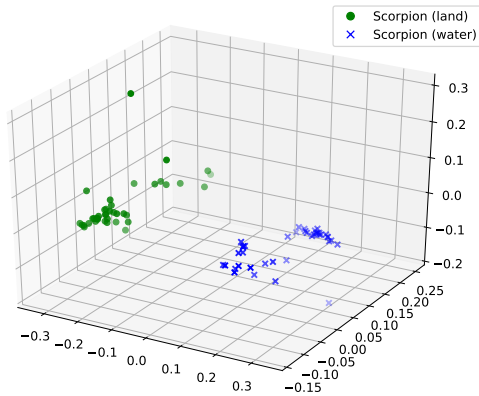
Figure 15: Low-dimensional MDS projections of the dissimilarity matrix of the Fastlizard creatures evolved in two different environments (on land or in water).

The body morphology was fixed within each of these populations, with the varying brain structures (where brain is understood as the neural network controlling the behavior of the creature) being a result of an evolutionary adaptation to the given environment.

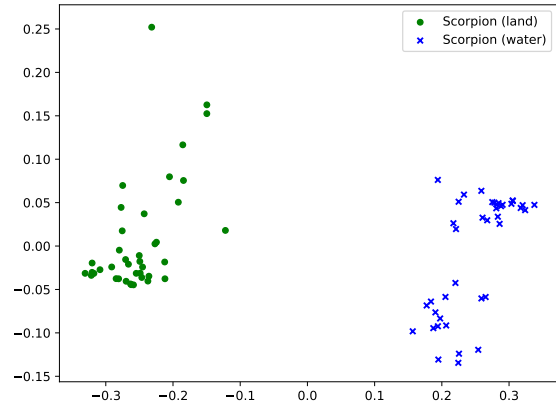
The comparison of the Fastlizard (land) and Fastlizard (water) creatures can be seen in Fig. 15. Both populations are clearly separated, with only one creature placed in the space between two, distinct clusters. Although both the specific weights and the equal weights provide a clear cut division between two populations, in the case of specific weights the creatures are clustered more tightly, with mean silhouette being 0.6 for specific weights and 0.37 for equal weights. This suggests that using the weights tailored towards the considered populations can help with differentiating between creatures evolved in different environments.

The comparison of the Scorpion (land) and Scorpion (water) creatures can be seen in Fig. 16. Just as in the case of the Fastlizard, both land and water creatures are separated by a wide margin. This observation is in agreement with the high values of the mean silhouette (0.77 for the specific weights, and 0.61 for the equal weights).

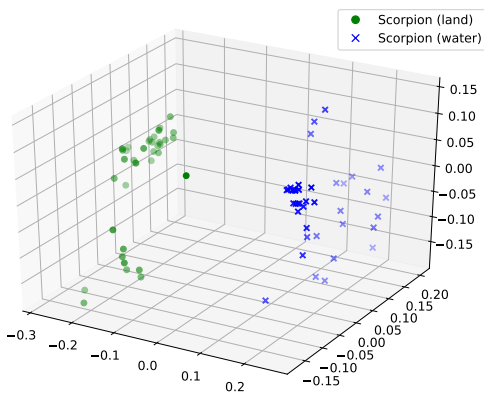
The comparison of the Star (land) and Star (water) creatures can be seen in Fig. 17. For these populations an important distinction can be made between the two-dimensional projections obtained for the specific and the equal weights: while the specific weights allow for a relatively clear distinction to be made between the parts of the space linked to specific environments, the use of the equal weights makes such a distinction impossible, as the two populations



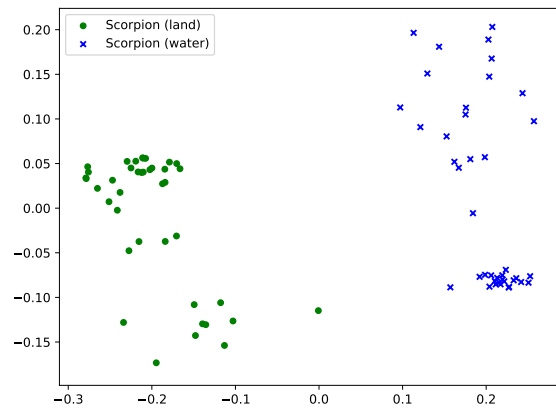
(a) Specific weights, 3D space (86.4% of the total variance preserved).



(b) Specific weights, 2D space (83.3% of the total variance preserved).



(c) Equal weights, 3D space (75.5% of the total variance preserved).

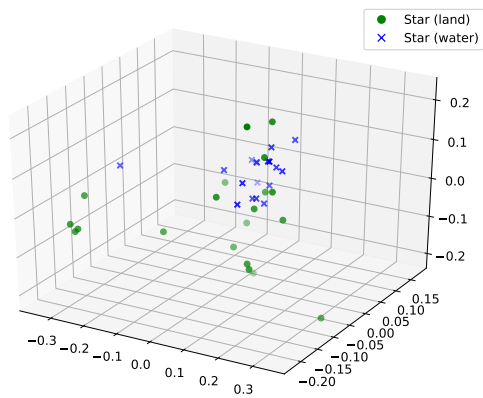


(d) Equal weights, 2D space (67.7% of the total variance preserved).

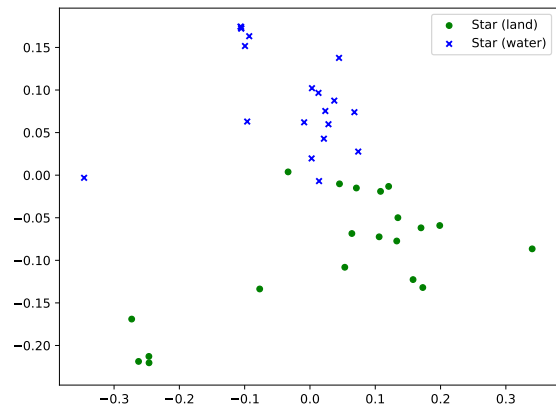
Figure 16: Low-dimensional MDS projections of the dissimilarity matrix of the Scorpion creatures evolved in two different environments (on land or in water).

overlap. The poor separation of the populations is confirmed by the low values of the mean silhouette, which are 0.25 for the specific weights, and only 0.13 for the equal weights. This suggests that for this type of creatures the properties of the movement do not strongly depend on the environment in which the creatures were evolved.

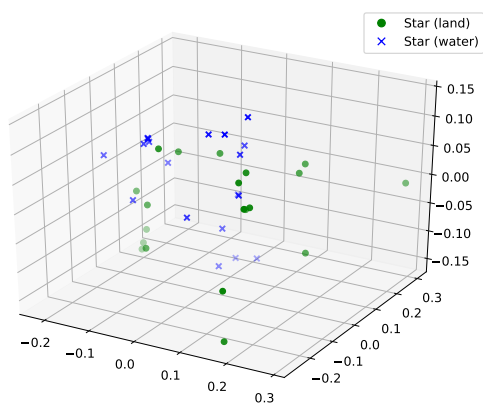
All the projections shown in Figs. 15, 16 and 17 preserve at least 45% of the variance present in the original space of eight dimensions, with some of them exceeding 80% of the preserved variation. Just as expected, the 3D projections preserve more of the variance than the 2D projections. More interesting is the fact that using the weights specific for the given populations results in a higher fraction of the variance being preserved than it is for the equal weights. The reason for that, is that although using the specific weights means that the variance of the already diversified properties is increased, the number of the high-variance properties is usually low, and so most of their variance can be included in the low-dimensional projections. Simultaneously the importance of the low-variance properties is accordingly decreased, and so the variance lost by not including them in the final projections is relatively smaller.



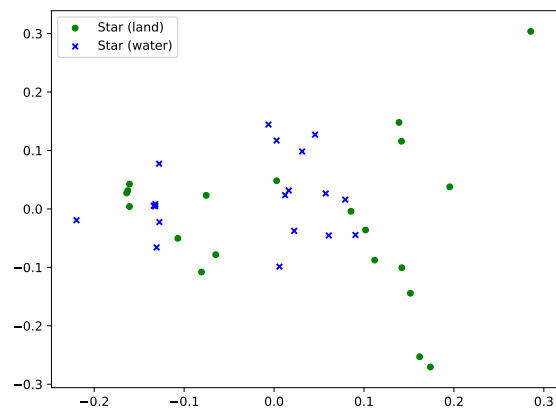
(a) Specific weights, 3D space (66.8% of the total variance preserved).



(b) Specific weights, 2D space (55.3% of the total variance preserved).

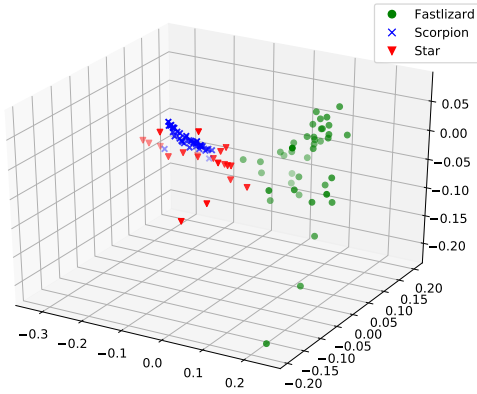


(c) Equal weights, 3D space (57% of the total variance preserved).

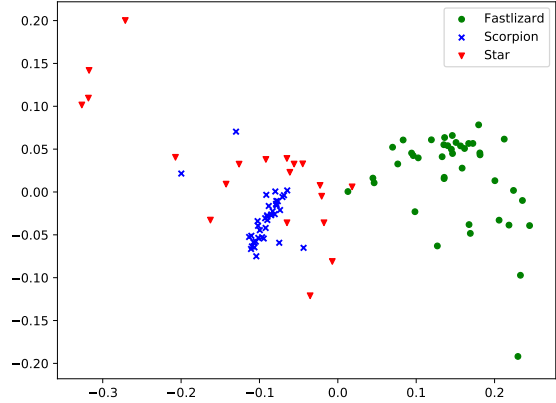


(d) Equal weights, 2D space (45.7% of the total variance preserved).

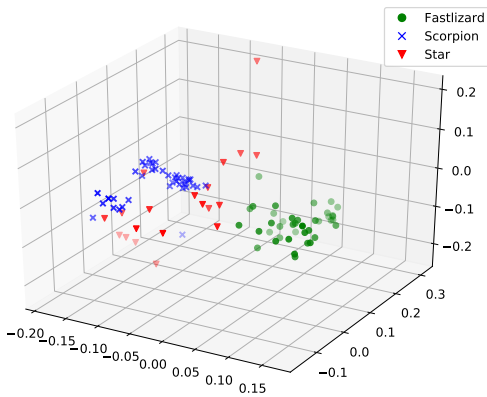
Figure 17: Low-dimensional MDS projections of the dissimilarity matrix of the Star creatures evolved in two different environments (on land or in water).



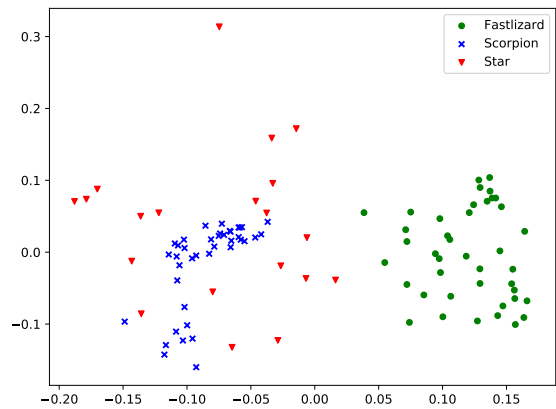
(a) Specific weights, 3D space (71% of the total variance preserved).



(b) Specific weights, 2D space (64.3% of the total variance preserved).



(c) Equal weights, 3D space (59.9% of the total variance preserved).



(d) Equal weights, 2D space (50% of the total variance preserved).

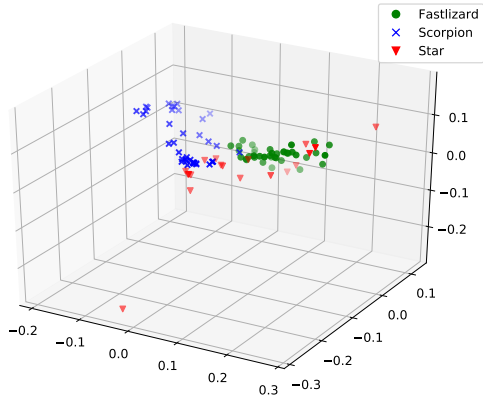
Figure 18: Land creatures dissimilarity matrices projected using MDS to low-dimensional spaces.

4.2 Estimating similarity between agents with reference to the test environment

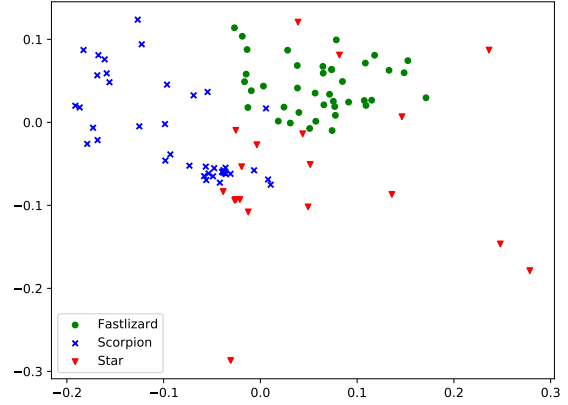
This experiment is analogous to the previous one, however instead of comparing the creatures with similar morphologies in different environments, the creatures with different morphologies evolved in the same environments are being compared.

First, three types of land creatures are being compared (Fig. 18). The three populations create three distinct clusters, with mean silhouette being 0.42 for the specific weights, and 0.37 for the equal weights. However, these clusters cannot be separated linearly quite as easily in the low-dimensional projections as it was the case for the water vs. land comparisons. This is true especially for the 2D projections, where Star creatures form a ring around a tight cluster of Scorpion creatures (nonetheless, these two populations do not appear to overlap).

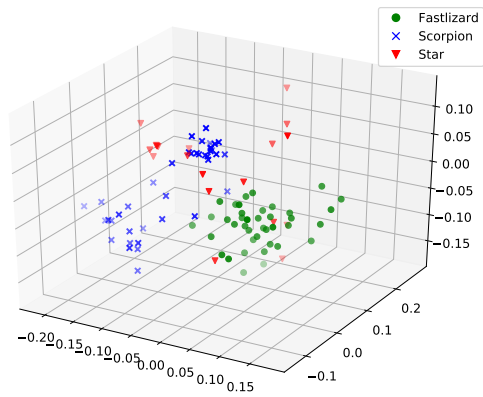
The results of the comparison between different types of creatures evolved in the water environment are presented in Fig. 19. The water creatures show a high variance which cannot be effectively captured by the low-dimensional projections – for the equal weights, the 2D projection captures only 37% of the variance present in the original space. This is a very low value, as even in the worst case scenario 25% of the variance would still be preserved. Although Fastlizard and Scorpion populations can be distinguished quite easily, the Star population – which shows the highest diversity among the considered populations – overlaps with the other



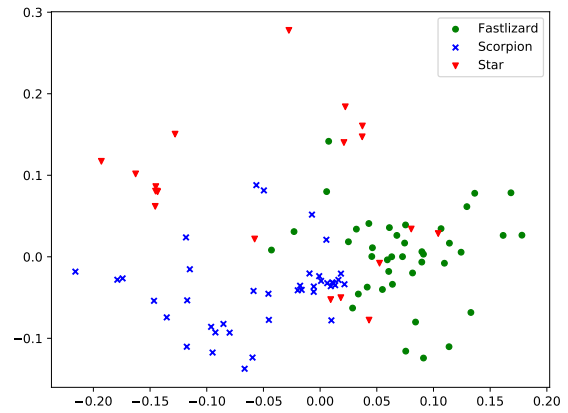
(a) Specific weights, 3D space (64.6% of the total variance preserved).



(b) Specific weights, 2D space (55.8% of the total variance preserved).



(c) Equal weights, 3D space (49.8% of the total variance preserved).



(d) Equal weights, 2D space (37.1% of the total variance preserved).

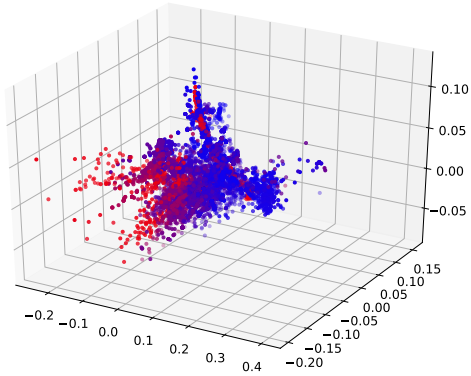
Figure 19: Water creatures dissimilarity matrices projected using MDS to low-dimensional spaces.

two, especially with the Fastlizard population. This overlap of populations can also be seen in the low values of the mean silhouette, which are 0.3 for the specific weights, and 0.22 for the equal weights.

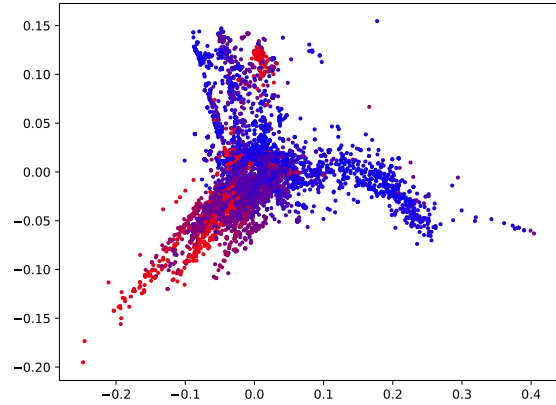
4.3 Estimating similarity between agents from many independent evolutionary runs

Due to the successful application of the customized weights in the previous section, only the population-specific weights of the properties are used in this section.

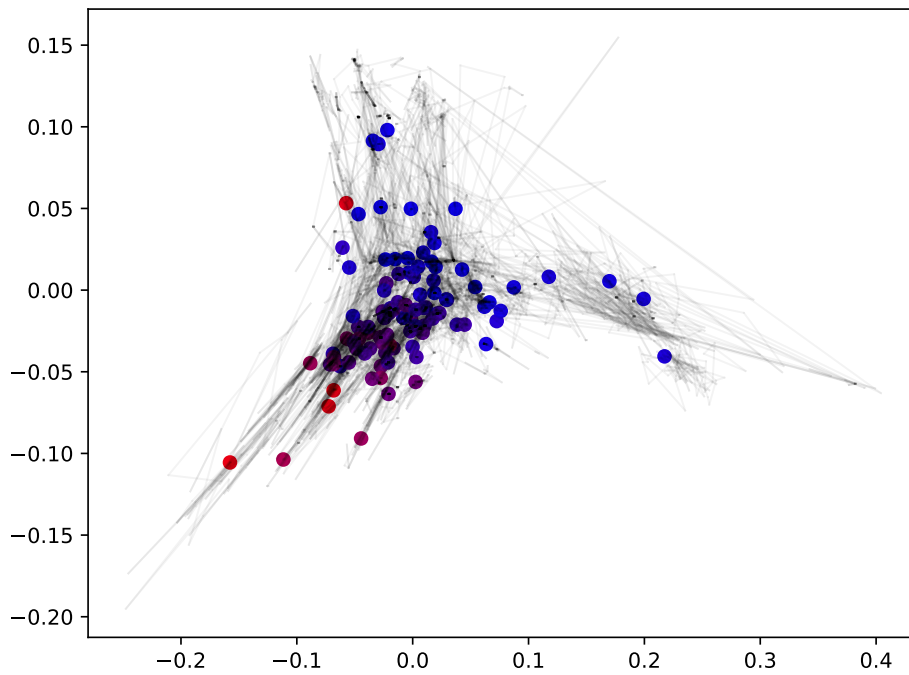
A multidimensional scaling of a big set of 10,000 creatures taken from 100 evolutionary runs (100 selected creatures each) was performed. Each of the selected creatures was chosen as the most fit creature among 1000 consecutively evaluated creatures in a given evolutionary run, with each evolutionary run evaluating 100,000 creatures in total. In all 100 evolutionary runs the creatures were evolved towards the goal of high speed, calculated as the distance covered by a creature over its lifespan divided by the length of its lifespan. The results of MDS are shown in Fig. 20. Less than a half of the creature variability has been covered by the low dimensional projections of the creature dissimilarity, which suggests that the evolution has explored a wide range of different behaviors. However, even with such a low amount of the preserved variability, some interesting structures can be observed in the data. The arrangement of the creatures in



(a) 3D projection of the creature dissimilarity (46.9% of the total variance preserved).



(b) 2D projection of the creature dissimilarity (38.4% of the total variance preserved).



(c) 2D projection of the creature dissimilarity with the trajectories of the evolutionary runs traced with black lines. The dots represent only the final creatures from each evolutionary run.

Figure 20: The dissimilarity of the creatures from one hundred independent evolutionary runs (100 creatures each, 10,000 in total). The color of each dot represents the fitness value of its corresponding creature (i.e., the mean speed over its lifespan), with blue representing the lowest and red representing the highest value.

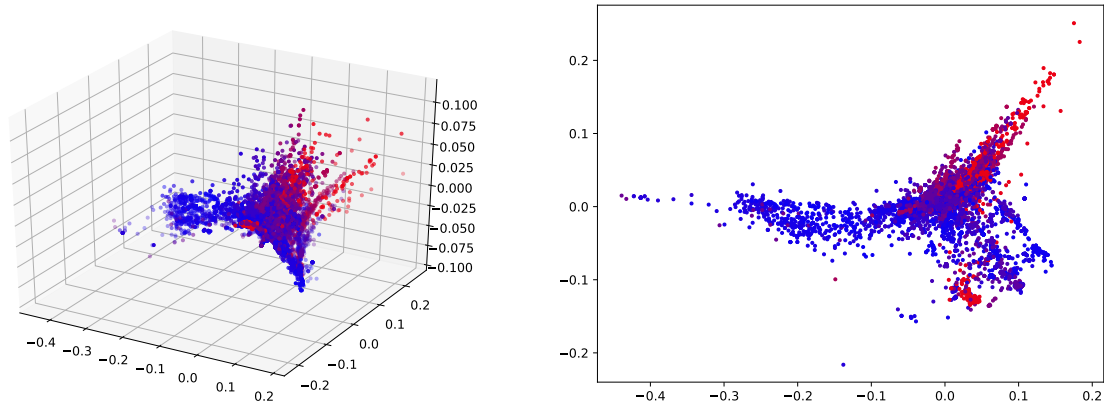
2D space resembles an outburst with three spikes protruding symmetrically on each side, facing away from a long, curved tail. These spikes may be composed of creatures following similar evolutionary paths, with each spike representing one path noticeably different from the others. The spikes facing downwards contain the fastest creatures, with the speed of the creatures growing towards the ends of the spikes: a gradient of the sought property (i.e., the high speed) can be observed in the space of the creature properties. This means that the proposed set of the properties could be used in the novelty search-based optimization of the creatures. However, one of the spikes facing upwards also contains fast creatures, suggesting that there is more than just one combination of the property values that yields fit creatures. Therefore the optimization problem of finding creatures with the highest speed is still multimodal in the space of movement properties.

As the gradients observed in Fig. 20 may result from a correlation between the speed of a creature, as measured over its whole life, and the mean instantaneous speed of that creature, the experiment has been repeated with *inst_speed* removed from the set of the properties. The results of this experiment are shown in Fig. 21. It appears, however, that the *inst_speed* was not responsible for the observed gradients, as they still can be seen even with that property ignored.

Another experiment explored a similar set of creatures taken from 10 independent evolutionary runs, 2000 selected creatures each. In order to reduce the computational load, the size of each evolutionary run sequence was reduced to 667 virtual creatures, with the properties of the virtual creatures being set as the median values of three consecutive creatures from the original sequence. The results of MDS are shown in Fig. 22. Similarly to the previous experiment, a general, imperfect gradient of the fitness values can be seen in the projections, with the speed of the creatures increasing from the right to the left side of the plots. In fact, even the cluster of fast creatures visible in the upper right corner of the plots comes from the evolutionary run that, eventually, ends up in the upper left corner of the plots.

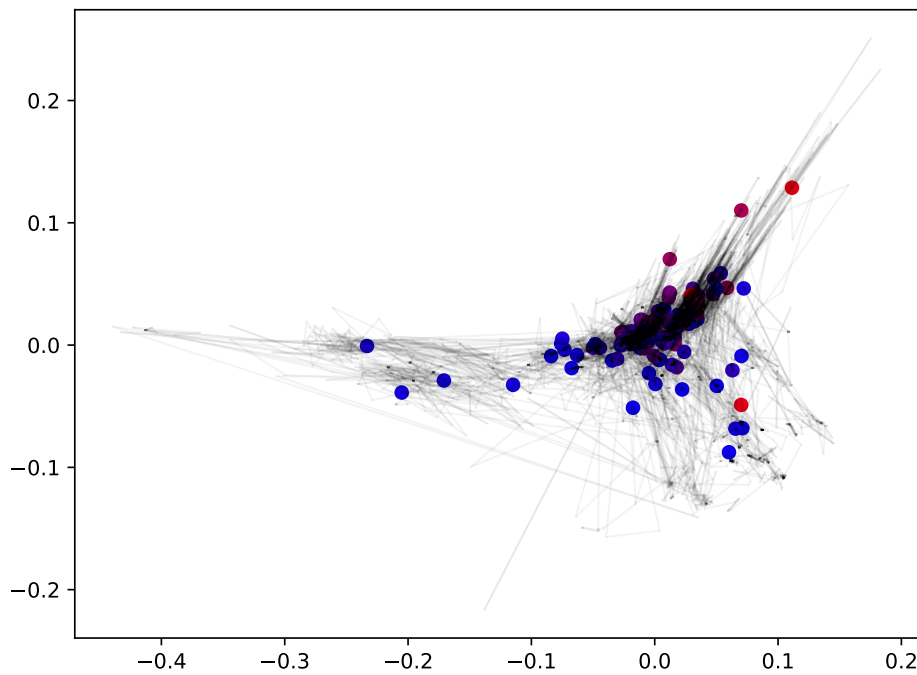
Some interesting observations can be made based on the shape of the trajectories taken by the separate evolutionary runs, as shown in Fig. 22(c). Although sometimes in the initial phase of the evolution the search process visits the unfit parts of the property space (e.g., the dark green trajectory), it eventually tends to end up in the more fit part of the property space on the left side of the plot. The exception from this rule is the black trajectory that appears to explore the parts of the property space that were not covered by any of the other evolutionary runs. It can also be observed that most of the trajectories end up in hard to escape local optima, with values of only some of the properties changing over time – this effect appears on the plots in the form of elongated clouds of uniform colors.

Just as before, additional experiments were performed with reduced sets of the properties in order to filter out the effects of those properties that were expected to most strongly correlate with the fitness values of the considered creatures. This time we have considered both excluding just *inst_speed*, and excluding *inst_speed* and *err_line_xy*. The results of these experiments are presented respectively in Figs. 23 and 24. The rationale behind excluding the *err_line_xy* property has been described in detail in Sect. 3.3. Although removing these properties from the set of the properties leads to some degree of an overlap between the creatures of different fitness levels, the gradient observed when the full set of the properties was used can still be seen. This means that no single property is responsible for the presence of that gradient in the projected data. It is worth noting however, that after excluding both *inst_speed* and *err_line_xy* from the set of considered properties, almost all of the final observations are clustered along one, elongated patch of the property space.



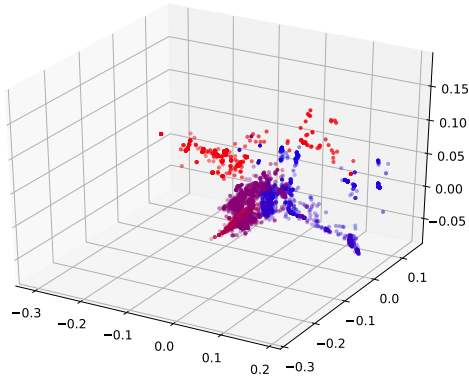
(a) 3D projection of the creature dissimilarity (58.2% of the total variance preserved).

(b) 2D projection of the creature dissimilarity (53.3% of the total variance preserved).

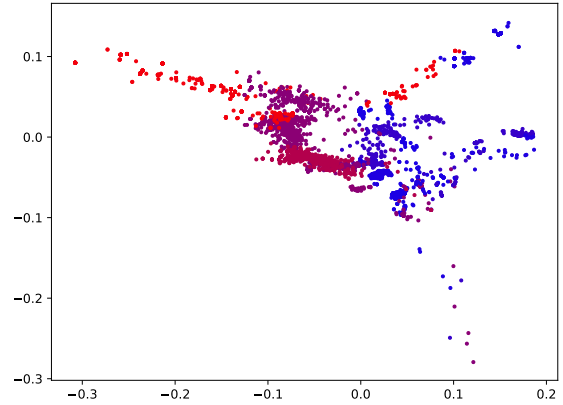


(c) 2D projection of the creature dissimilarity with the trajectories of the evolutionary runs traced with black lines. The dots represent only the final creatures from each evolutionary run.

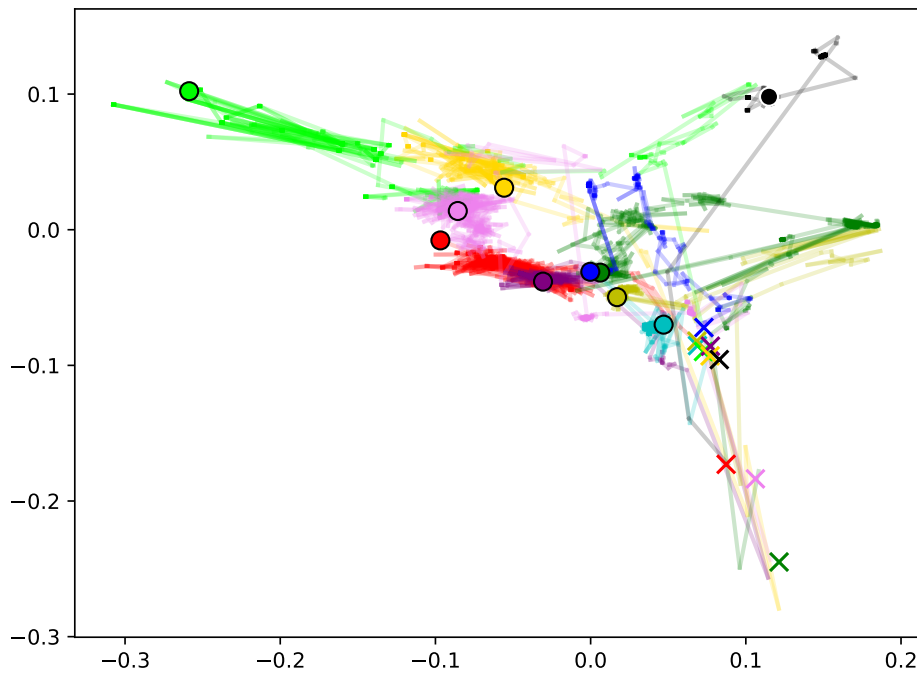
Figure 21: The dissimilarity of the creatures from one hundred independent evolutionary runs (100 creatures each, 10,000 in total). The color of each dot represents the fitness value of its corresponding creature (i.e., the mean speed over its lifespan), with blue representing the lowest and red representing the highest value. *inst_speed* has been excluded from the set of the considered properties in order to remove the expected influence of its correlation with the fitness values of the creatures.



(a) 3D projection of the creature dissimilarity (72% of the total variance preserved).

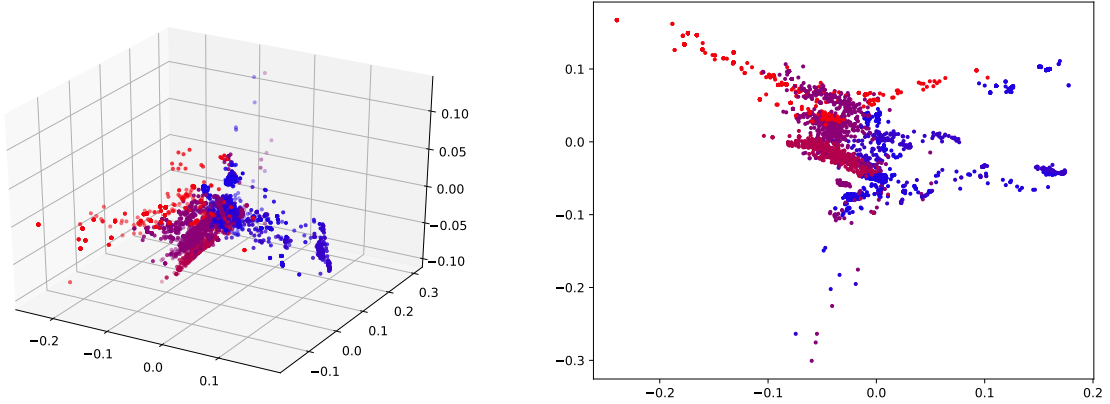


(b) 2D projection of the creature dissimilarity (67% of the total variance preserved).



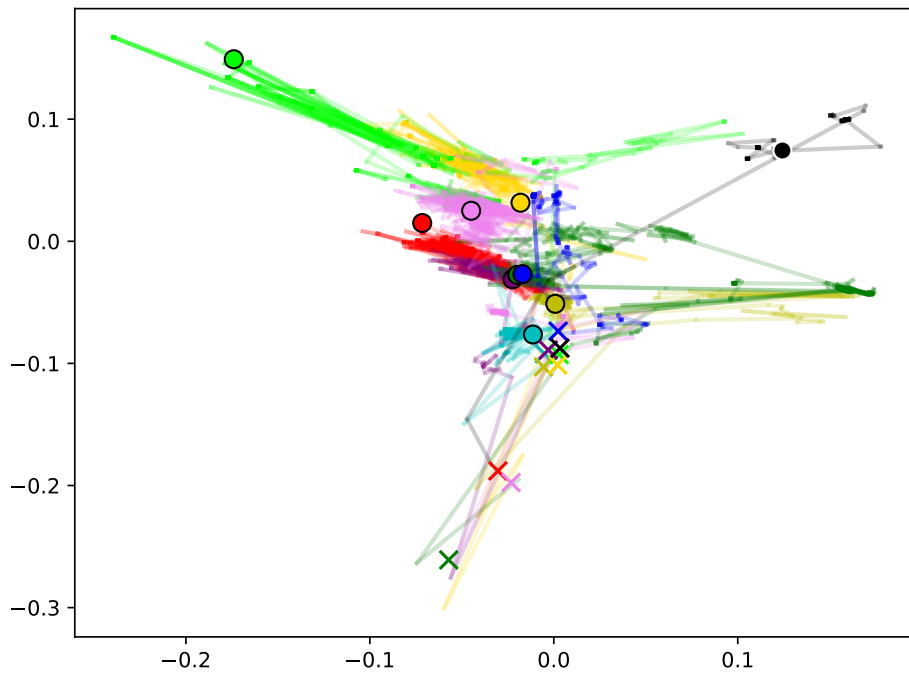
(c) 2D projection of the creature dissimilarity with the trajectories of the evolutionary runs traced with lines (each color corresponds to one run). The crosses represent the starting points, and the dots represent the ending points of each evolutionary run.

Figure 22: The dissimilarity of the creatures from ten independent evolutionary runs (1000 creatures each, 10,000 in total). In Figs. 22(a) and 22(b), the color of each dot represents the fitness value of its corresponding creature (i.e., the mean speed over its lifespan), with blue representing the lowest and red representing the highest value.



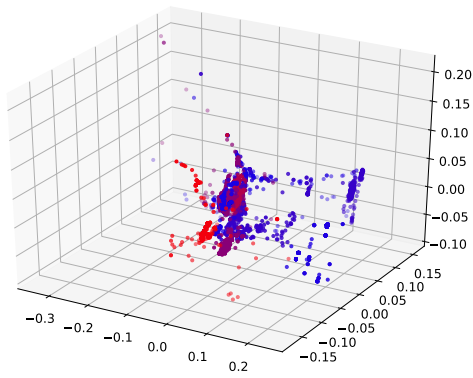
(a) 3D projection of the creature dissimilarity (68.8% of the total variance preserved).

(b) 2D projection of the creature dissimilarity (63.7% of the total variance preserved).

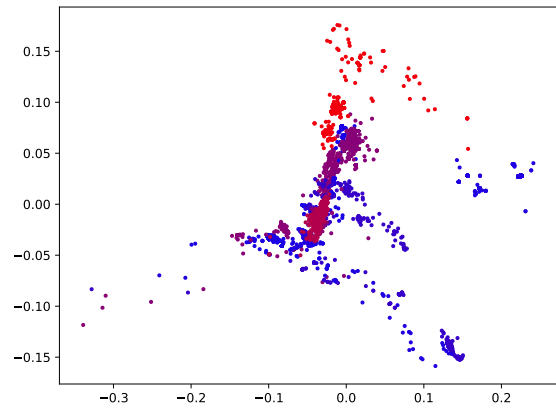


(c) 2D projection of the creature dissimilarity with the trajectories of the evolutionary runs traced with lines (each color corresponds to one run). The crosses represent the starting points, and the dots represent the ending points of each evolutionary run.

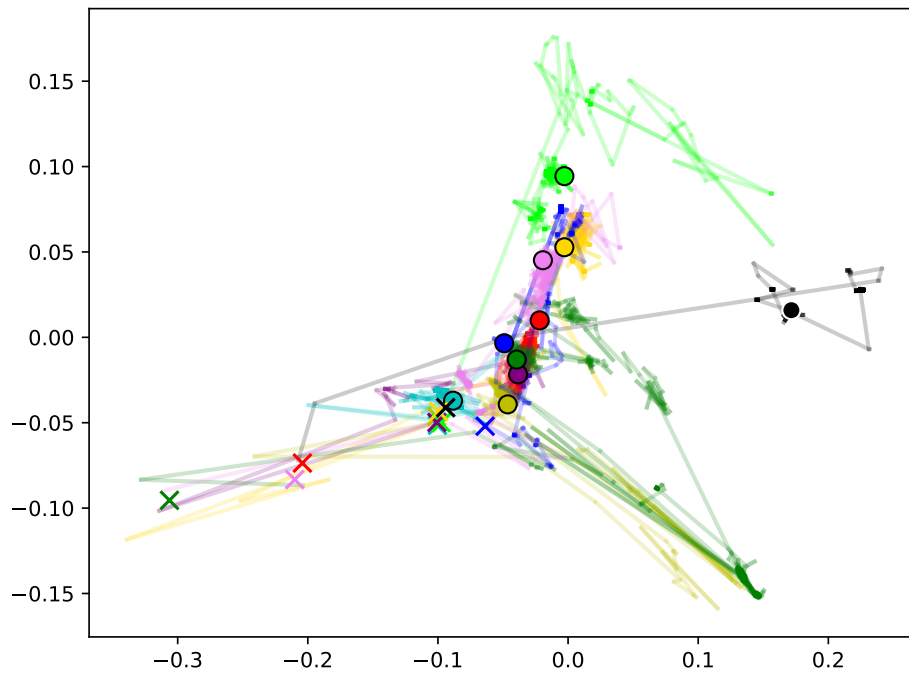
Figure 23: The dissimilarity of the creatures from ten independent evolutionary runs (1000 creatures each, 10,000 in total). In Figs. 23(a) and 23(b), the color of each dot represents the fitness value of its corresponding creature (i.e., the mean speed over its lifespan), with blue representing the lowest and red representing the highest value. *inst_speed* has been excluded from the set of the considered properties in order to remove the expected influence of its correlation with the fitness values of the creatures.



(a) 3D projection of the creature dissimilarity (71.4% of the total variance preserved).



(b) 2D projection of the creature dissimilarity (66.2% of the total variance preserved).



(c) 2D projection of the creature dissimilarity with the trajectories of the evolutionary runs traced with lines (each color corresponds to one run). The crosses represent the starting points, and the dots represent the ending points of each evolutionary run.

Figure 24: The dissimilarity of the creatures from ten independent evolutionary runs (1000 creatures each, 10,000 in total). In Figs. 24(a) and 24(b), the color of each dot represents the fitness value of its corresponding creature (i.e., the mean speed over its lifespan), with blue representing the lowest and red representing the highest value. *inst_speed* and *err_line_xy* have been excluded from the set of the considered properties in order to remove the expected influence of their correlation with the fitness values of the creatures.

5 Summary and further work

This report introduced a number of basic properties of movement that can be easily estimated for real or simulated creatures, or robots consisting of moving parts consistent with the model described in Sect. 2. In Sect. 3 we focused on technical description of each property that captures a single aspect of movement, as well as on demonstrating and verifying their characteristics. Sect. 4 analyzed a few populations of evolved creatures using properties of movement, demonstrating that different environments or conditions under which the creatures are evolved can be captured by the introduced properties of movement. A more comprehensive analysis of a larger set of evolved creatures is available [11].

In order to verify the characteristics of each properties of movement introduced in this work, more experiments would be useful. Such experiments include the analysis of creatures evolved in different conditions (gravity, slipperiness, water depth), assigning human labels to clusters of movement methods discovered by the movement properties, automatically identifying representative agents for each cluster [6], discovering weights or aggregation methods of movement properties (possibly with machine learning models) that reflect human classification of movement strategies, and investigating and visualizing how average and individual movement strategies change during evolution. Additional properties that would expand the proposed set of the properties could also be considered, including the properties based on the mean speed of a creature as a function of a length of the time window in which that speed is being calculated. Scalability of all the properties (i.e., invariance of the values of properties to varying resolution of sampling of the points on the body of a creature) should be examined, and the properties that lack this quality should be improved.

References

- [1] A. D. Aczel. *Complete Business Statistics*. Tata McGraw-Hill, 2006.
- [2] David R. Brillinger. *Time Series: Data Analysis and Theory*. SIAM, 2001.
- [3] F. Castella, F. Happé, U. Fritha, and C. Frithc. Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage*, 12:314–325, 2000.
- [4] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, USA, third edition, 1996.
- [5] K. Grill-Spector. The neural basis of object perception. *Current Opinion in Neurobiology*, 13:159–166, 2003.
- [6] Maciej Komosinski. Applications of a similarity measure in the analysis of populations of 3D agents. *Journal of Computational Science*, 21:407–418, 2017. URL: <http://dx.doi.org/10.1016/j.jocs.2016.10.004>, doi:10.1016/j.jocs.2016.10.004.
- [7] Maciej Komosinski, Grzegorz Koczyk, and Marek Kubiak. On estimating similarity of artificial and real organisms. *Theory in Biosciences*, 120(3-4):271–286, December 2001. URL: <http://dx.doi.org/10.1007/s12064-001-0023-y>, doi:10.1007/s12064-001-0023-y.
- [8] Maciej Komosinski and Marek Kubiak. Quantitative measure of structural and geometric similarity of 3D morphologies. *Complexity*, 16(6):40–52, 2011. URL: <http://dx.doi.org/10.1002/cplx.20367>, doi:10.1002/cplx.20367.
- [9] Maciej Komosinski and Agnieszka Mensfelt. A flexible dissimilarity measure for active and passive 3D structures and its application in the fitness–distance analysis. In

- Paul Kaufmann and Pedro A. Castillo, editors, *Applications of Evolutionary Computation*, pages 106–121, Cham, 2019. Springer. URL: <http://www.framsticks.com/files/common/DissimilarityMeasure3DStructuresFitnessDistance.pdf>, doi: 10.1007/978-3-030-16692-2_8.
- [10] Maciej Komosinski, Agnieszka Mensfelt, Paweł Topa, and Jarosław Tyszk. Application of a morphological similarity measure to the analysis of shell morphogenesis in Foraminifera. In Aleksandra Gruca, Agnieszka Brachman, Stanisław Kozielski, and Tadeusz Czachórski, editors, *Man–Machine Interactions 4*, volume 391 of *Advances in Intelligent Systems and Computing*, pages 215–224. Springer, 2016. URL: <http://www.framsticks.com/files/common/ForaminiferaGenotypePhenotypeMapping.pdf>, doi: 10.1007/978-3-319-23437-3_18.
- [11] Maciej Komosinski and Konrad Miazga. Measuring properties of movement in populations of evolved 3D agents. In Harold Fellermann, Jaume Bacardit, Ángel Goñi-Moreno, and Rudolf M. Füchslin, editors, *Artificial Life Conference Proceedings*, pages 485–492. MIT Press, 2019. doi:10.1162/isal_a_00208.
- [12] Maciej Komosinski and Krzysztof Rosinski. Estimating similarity of neural network dynamics. Research report RA–10/10, Poznan University of Technology, Institute of Computing Science, 2010. URL: <http://www.framsticks.com/files/common/SimilarityNeuralNetworkDynamics.pdf>.
- [13] Maciej Komosinski and Szymon Ulatowski. Framsticks SDK (Software Development Kit). URL: <http://www.framsticks.com/sdk>.
- [14] Maciej Komosinski and Szymon Ulatowski. Genetic mappings in artificial genomes. *Theory in Biosciences*, 123(2):125–137, September 2004. URL: <http://dx.doi.org/10.1016/j.thbio.2004.04.002>, doi:10.1016/j.thbio.2004.04.002.
- [15] Maciej Komosinski and Szymon Ulatowski. Framsticks web site, 2019. URL: <http://www.framsticks.com>.
- [16] Caroline Larboulette and Sylvie Gibet. A review of computable expressive descriptors of human motion. In *Proceedings of the 2nd International Workshop on Movement and Computing*, MOCO '15, pages 21–28, New York, NY, USA, 2015. ACM. URL: http://zabador.free.fr/Publications/2015/LG15/moco15_larboulette_gibet.pdf, doi: 10.1145/2790994.2790998.
- [17] J. M. Bland and S. M. Kerry. Weighted comparison of means. *British Medical Journal*, 316, 1998.
- [18] W. Mendenhall and T. Sincich. *A Second Course in Statistics: Regression Analysis*. Prentice Hall, 1996.
- [19] Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics (SIAM), 2000.
- [20] B. Noble and J. Daniel. *Applied Linear Algebra*. Prentice Hall, 1988.
- [21] M. V. Peelen and P. E. Downing. The neural basis of visual body perception. *Neuroscience*, 8:636–648, 2007.
- [22] David Stephen Pollock. *A Handbook of Time-Series Analysis, Signal Processing and Dynamics*. Academic Press, 1999.

- [23] J.A. Pyles, J.O. Garcia, D.D. Hoffman, and E.D. Grossman. Visual perception and neural correlates of novel ‘biological motion’. *Vision Research*, 47(21):2786–2797, 2007. URL: <http://dx.doi.org/10.1016/j.visres.2007.07.017>, doi:10.1016/j.visres.2007.07.017.
- [24] J.A. Pyles and E.D. Grossman. Neural adaptation for novel objects during dynamic articulation. *Neuropsychologia*, 47(5):1261–1268, 2009.
- [25] Peter Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, 20(1):53–65, November 1987. doi:10.1016/0377-0427(87)90125-7.
- [26] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas. The earth mover’s distance as a metric for image retrieval. *International journal of computer vision*, 40(2):99–121, 2000.
- [27] Yaccov (J) Stein. Two dimensional Euclidean regression. 1983. URL: <http://www.dspscsp.com/pubs/euclreg.pdf>.