14

Gestural Expression

Dans le vide de l'espace quelqu'un dessine,
Crée à travers son corps l'infini du temps.
Marcel Marceau

Gestural expression is a difficult topic, not because the concept of a gesture is so difficult, but because it looks so obvious that everybody believes to know what it means. This first impression is, however, misleading and it is there that the difficulty arises. The word "gesture" is like "time": If your are not asked, what it means, you know, but if you are asked you cannot tell. Saint Augustin's famously articulated this fact when asked about time. We shall not give a precise definition of a gesture until section 14.8. Instead, we want to work with an intuitive understanding.

14.1 General Facts About Gesture Theory in Music

Despite the consciousness that gestures play an important role in music, there has not been such a thing like a gesture theory in music. This was made explicit by music theorist Robert S. Hatten in his book *Interpreting Musial Gestures, Topics, and Tropes* [47]: "Given the importance of gesture to interpretation, why do we not have a comprehensive theory of gesture in music?" Gesturality became an important topic when Hatten realized that performance of classical piano music—Mozart, Beethoven, Schubert—is strongly determined by gestural utterance. Although Hatten's investigations are focusing on gesture, they are not formalized but remarkably subtle. His definition of a gesture reads as follows: "Gesture is most generally defined as communicative (whether intended or not), expressive, energetic shaping through time (including characteristic features of musicality such as beat, rhythm, timing of exchanges,

G. Mazzola, *Musical Performance*, Computational Music Science,
DOI 10.1007/978-3-642-11838-8_14, © Springer-Verlag Berlin Heidelberg 2011

contour, intensity), regardless of medium (channel) or sensory-motor source (intermodal or cross-modal)."

On a more formal level of music theory, David Lewin, the *monstre sacreé* of American music theory, has indicated that lack of gesturality in his seminal book *Generalized Musical Intervals* [68]. Lewin suggests that transformations between musical points (pitch classes, for example) are the new path to pursue. In [68, p. 159], we read: "If I am at s and wish to get to t, what characteristic gesture should I perform in order to arrive there?" Although his transformational theory is in fact not a gestural one yet (it is merely the step into mathematical—more precisely, category-theoretical—process diagrams), Lewin thinks in a gestural spirit and repeats this attitude in another statement, still relating to his question about the movement of s to t. He adds [68, p. 159]: "This attitude is by and large the attitude of someone *inside* the music, as idealized dancer and/or singer. No external observer (analyst, listener) is needed."

Interestingly, Hatten's and Lewin's approaches are based upon a quite blurred concept of a gesture, but they have the merit to have made that important point. It is not only in academic music theory that this insight has been forwarded. Piano giant and music philosopher Cecil Taylor also pointed out, "The body is in no way supposed to get involved in Western music. I try to imitate on the piano the leaps in space a dancer makes."

So the gestural dimension is recognized, but not elaborate, on a theoretical level. This has strong impliciations for performance theory of gestures: Everybody agrees that performance of gestural contents and performance via rhetorical application of gestures is crucial, but when it comes to the explicit, detailed discourse, the topic evaporates like that beautiful, but equally ghostly, discourse about playing the body of time. Although a number of gesture-driven investigations and implementations have been realized by computer music communities [146], these contributions do not improve the conceptual control of the topic because they are restricted to applying the physical gestures as spatial movements of the body's limbs to specialized interfaces with electronic musical instruments.

14.2 Roger Sessions: Gestures in Performance

The dramatically intense but still underestimated role of gestures in performance has been described in a beautifully clear way by American composer and music critic Roger Sessions in his book *Questions About Music* [125, Chapter III]:

> *It is fairly obvious, I suppose, that our total awareness of movement—which in essence signifies our awareness of* time as a process*—demands sustained attention, which is limited to the duration of the specific act*

of movement in question; it holds us captive, as it were, for the duration. We are aware of a beginning and an end. In respect to space on the other hand, the words "beginning" and "end" have an essentially metaphorical meaning; they represent boundaries or limits that remain even after we have become aware of them, as does all that lies between. Our attention is our own to husband and deploy as we wish. We can withdraw it and absent ourselves merely by averting or closing our eyes, and return whenever and for as long as we wish.

What I am saying is that we experience music as a pattern of movement, as a gesture; and that a gesture gradually loses its meaning for us insofar as we become aware of having witnessed it, in its total identity, before. If it is to retain this meaning in its full force, it must be on each occasion reinvested with fresh energy. Otherwise we experience it, to an increasing degree, as static; its impact, as movement, diminishes, and in the end we cease to experience it as movement at all. Its essentially static nature has imposed itself on our awareness.

This is why I am convinced that the performer is an essential element in the whole musical picture. It is why I came to realize that my earlier dreams—that composers might learn to freeze their own performance, in wax or otherwise (tape recorders had not been invented at that time)—were, to put it bluntly, quite ill-directed. They were ill-directed, above all, for the reasons I have been outlining; a gesture needs constant renewal if it is to retain its force on subsequent repetitions. Composers above all should know this, especially if they have developed the practice of taking part in performances of thir own work. Each performance is a new one, and the work is always studied and approached anew, even by the composer. The same, it should be obvious, is true of professional performers. I would go even much further and point out that there is no such thing as a "definitive" performance of any work whatever. This is true even of performances by the composer himself, in spite of the fact that recordings of his performances of his own work should be made and preserved, for a number of quite obvious reasons.

Session's discussion of movement as a processing of time leads him to acknowledge that this dynamical action is a gesture—not only in the making, but also in the music's perception. So he gives the argument for a messaging of gestures, and by means of gestures, which is our topic in this chapter. It is remarkable that he then recognizes that a gesture cannot preserve its meaning except in its energetic refreshment on each occasion of performance. This is very similar to the French theory of gestures [91, Chapter 7.2], which stresses the impossibility to tame living gestures.

He moreover recognizes the performer's essential role in the "whole music picture" and also reminds composers, himself included, that their work of musical creation is not accomplished until it is performed. This does not mean that a composer must intervene in the performance of his/her works. Some are

dead and simply cannot do this anymore. No, it means that the completion of a musical work cannot be achieved before its performance has occured. In this sense, performance is strongly what semioticians call a *deictic* part of the musical sign system: Musical signs reach their full meaning only and essentially through their pragmatic instantiation.

Fig. 14.1. Roger Sessions.

This second insight is strongly related to the gestural aspect since gestures are not lexicographic, they are shifters, as Sessions stresses with his "French" view on gestures. We are not astonished that Manfred Clynes refers to Session's writings in his critique of score-based music. We come back to the gestural aspect in Clynes' work later in section 14.5.

14.3 Theodor Wiesengrund Adorno's Gesture Theory in Musical Reproduction

Theodor W. Adorno has written deep analyses of performance, in particular with respect to their subcutaneous gestural implications in his posthumous work *Zu einer Theorie der musikalischen Reproduktion* [2]. It is interesting to see how Adorno gets off the ground with his gestural discourse on the same basis as Sessions and Clynes, namely a radical critique of the score-based reduction of music (translated from [2, p.227/8]):

> *Notation wants music to be forgotten, in order to fix it and to cast it into identical reproduction, namely the objectivation of the gesture, which for all music of barbarian cultures martyrs the eardrum of the listener. The eternization of music through notation contains a deadly moment: what it captures becomes irrevocable.*
>
> (...)
>
> *Spatialization (through notation) means total control. This is the utopic contradiction in the reproduction of music: to re-create by total control what had been irrevocably lost.*
>
> (...)
>
> *All making music is a recherche du temps perdu.*

And later on (translated from [2, p.235]):

> *Musical notation is an expression of the Christianization of music.*
>
> (...)
>
> *It is about eternity: it kills music as a natural phenomenon in order to conserve (or "embalm" G.M.) it—once it is broken—as a spiritual entity: The survival of music in its persistence presupposes the killing*

of its here and now, and achieves within the notation the ban (or "detachment" G.M.) from its mimetic[1] *representation.*

To begin with, Adorno, Sessions, and Clynes agree upon the fact that music notation, and its score, abolishes music, which is fixed and cast into a format for identical reproduction. It does so in objectifying the gesture and thereby martyring the eardrum, an act of barbarian culture. It is remarkable that musical notation is related to barbarian culture. The eternization of music in the notation's casting is killing music; it retains a dead body, not the living music. This eternity of dead—in fact, embalmed—bodies appears as a Christian ritual of sacred denaturation. The procedure of notation kills the music's here and now; its expressivity is annihilated, banned forever. The notational process kills through spatialization, which means total control, time does not fly by anymore, a note is a point in a dead space of eternity. Adorno views this as being the great contradiction of notation in that it claims total control for a reproduction of what has been irrevocably lost. It is a *temps perdu*, and making music is doomed to a *recherche du temps perdu*.

Adorno then makes important comments on what he views as being the gestural substance of music (translated from [2, p.244/5]):

As each face and each gesture, each play of features, is mediated by the I, so the musical moments are the very arena of mimic in music. What must be read and decoded within music are its mimic innervations.

(...)

However, a pathetic or cautious or expiring location does not signify pathos, caution, or expiration as a spiritual thing, but maps the corresponding expressive categories into the musical configuration, and those who want to perform them correctly have to find those encapsulated gestures in order to mimic them.

(...)

Finding through reading: the decoding work by the interpreter; the very concept of musical performance is the path into the empire of mimic characters.

(...)

The spatialization of gestures, that impulse of neumatic notation is at the same time the negation of the gestural element.

(...)

By the visual fixation, where the musical gesture is positioned into a simultaneous relation to its equals, it ceases to be a gesture, it becomes an object, a mental thing.

Here Adorno refers to the mimetic category in his theory. It is the category "expression of expression." So it is about the expression of emotions,

[1] For Adorno, "mimesis" means "expression of expression," and this is precisely our context: The expression as content is expressed via rhetorical shaping.

for example, not about emotions, and it is about the musical image of these expressions. Therefore, we have to read those mimic innervations of gestural expressions in music. Musical performance deals with the explication of those hidden innervations, with the action of displaying them in the making, here and now. And it now becomes clear that the neumatic notation creates static photographs of those gestures, which negate them by this spatial fixation. The spatial trace of a gesture is its negation, freezing it as a spatial object.

We should, however, briefly digress on the very concept of a space here, since it is not what a geometer or a physicist would call a space. In physics, a space is a geometric entity that can have different interpretations, so space-time is (locally speaking) a four-dimensional real vector space, and the mathematical structure of time is not different from that of the three space coordinates. Of course, the Lorentz metric distinguishes time in the metrical structure of space-time, but it is still a metrical space. In performance theory of music, time has a radically different role. The four-dimensional space of onset, pitch, loudness, and duration for piano music, which is used in score notation, does not have the ontology of musical time. Under no circumstances would the onset or duration coordinates be accepted as representing the time that takes place in performance. This *differentia specifica* in the performative time concept is related to gestures, not to geometric representation. For Adorno, gesture has an existential character; it cannot be objectivized; it only exists in the moment of the making; it is mediated by the I, which cannot be cast in a dictionary—the I is the non-lexical, the shifter, par excellence. However, it is not part of the subject, it is not subjective as opposed to being objective (the score objects are so). I is only mediated by the I, it seems to lie between subject and object; therefore, the utterance of a gesture is neither object nor subject.

Adorno continues (translated from [2, p.269]):

> *The true reproduction is the mimicry of a non-existent original.*
> (...)
> *But this mimicry of the non-existent original is at the same time nothing else but the X-ray photograpy of the text.*
> (...)
> *Its challenge is to make evident all relations, transitions, contrasts, tension and relaxation fields, and whatever there is that builds the construction, all of that being hidden under the mensural notation and the sensorial surface of sounds.*

The true reproduction is not a reference to an object out there; the original is non-existent, and it is not the I, which would be an existent entity. It is something mysterious since there is an X-ray procedure, but it does not show something hidden in the dead object of the score. It is as if that mystery would be brought to existence by the very X-ray procedure. The innervation must be made, not only discovered and pointed to.

Adorno's concept of a gesture is as difficult as it is radically different from what can be described in terms of traditional subject-object duality.

Let us see what Adorno concludes from all these subtle reflections (translated from [2, p.269,270,271]):

> *What happens in true performance is the articulation of the sensorial appearance that reaches into the most hidden details, wherein the totality of the construction, the gesture of the work, reveals its mimical execution.*
>
> (...)
>
> *The concept of clarity defines the degree of an analytical performance: everything that exists as relations within the mensural text must become clear, but this concept cannot be understood in a primitive way, i.e. as a clarity of every single relation, but as a hierarchy of clarity and blurredness in the sense of the clarity of the overall structure, the mimic gesture.*

And he summarizes this entire perspective on gestural performance (translated from [2, p.247]):

> *Correspondingly the task of the interpreter would be to consider the notes until they are transformed into original manuscripts under the insistent eye of the observer; however not as images of the author's emotion—they are also such, but only accidentally—but as the seismographic curves, which the body has left to the music in its gestural vibrations.*

14.4 Renate Wieland's Gestural Piano Pedagogy

As a student of Adorno, piano pedagogue Renate Wieland, in collaboration with her colleague Jürgen Uhde, has developed a theory of piano performance that is based upon Adorno's gestural philosophy. The remarkable feature of this work is that she succeeds in

(1) giving her approach a clear-cut separation from emotional dramaturgy and

(2) reshaping gesture theory in an explicit geometric language.

Fig. 14.2. Renate Wieland.

She makes these two points very clear in her text (translated from [140, p.169]):

> *Musical gestures are perceived in the free conducting movement, in the playing movement and sublimated in the spiritual mimesis of pure imagination. Whatever the level, such experiments are always within space. Originally, affects were actions, related to an exterior object,*

along the process of interiorization they were detached from their object, but they are still determined by the coordinates of space.

(...)

Language reminds us everywhere of the connection of affect and movement and of the way gestures behave in space. It speaks about hautiness, elevation and inclination, about greatness of mind, pettiness, about respectful and forward, etc.

(...)

There is therefore something like gestural coordinates; they can help ask how the gestural impulse out of the inner is projected into space, how it wants to expand, which direction is dominant: Is its energy vertically or horizontally active? Does it rather propagate ahead or backward? Upward or downward? To the right or to the left? Are forces acting more concentrically or excentrically? Does the gesture rather point "inward," as we read in Schumann's work, or "outward"? Which amplitude does the expression choose? Does it live in all spatial dimensions, and with what proportion and intensity?

She reminds us of the etymology of the word "emotion": *ex movere*, to move from inside out. She also makes clear that the original setup is now internalized, but that it remains a spatial concept. She then gives examples of etymological shifts, which are parallel to this internalization process: Words now mean abstract things, but when we go to the kernel of that meaning, it is related to a spatial action. So the mimetic action in Adorno's sense is the expression of that spatially conceived gesture in the realm of musical space.

She adds the following excellent illustration of a gestural mimesis in music (translated from [140, p.169]):

Models of contrast between extreme vertical and horizontal gestures are found in Beethoven's Bagatelle op.126,2.

(...)

Aggressively starting initial gestures are answered by flat, conciliating gestures, where the extremes are polarized to the outermost in the course of the piece. In this way, asking again and again, gesture becomes plastic in the end. But it only succeeds insofar as it constitutes a unity, is emanated from one inner central impulse.

(...)

Gestures are the utmost delicate; where their unity is disturbed, their expression immediately vanishes.

It is again in Adorno's and Session's spirit that she views gestures as being extremely unstable in their existentiality: Nothing is easier than to disturb and vaporize a gesture. It is by this fact that Gilles Châtelet, one of the fathers of French gesture theory, has characterized gestures as being the smile of existence [14].

Wieland finally transcends her approach in a seemingly breathtaking intensification, which reads as follows (translated from [140, p.190]):

The touch of sound is the target of the comprising gesture; the touch is so-to-speak the gesture within the gesture, and like the gesture at large, it equally relates to the coordinates of space.

(...)

The eros of the pianist's touch is not limited to the direct contact with the key, the inner surface of the entire hand pre-senses the sound. etc. etc.

She introduces what one could call the reverberance of a gesture, namely the gesture within a gesture, meaning that a gesture can incorporate other gestures, can become a gesture of gestures. We shall see later in section 14.8, relating to our own research, that this concept is very powerful for the theory of gestures in that it enables complex imbrications of gestures, so-called hypergestures, for the construction of movements of movements of movements..., an idea that is crucial in the dynamics of musical utterances.

14.5 Manfred Clynes' Essentics as a Theory of Gestural Expressivity

We saw in section 13.3 that Manred Clynes conceived expressivity as a shaping of performance in pulses, those embodiments of essentic forms, via specific deformations of duration and loudness. He claimed that such pulses were characteristic for the emotional expressivity of composers such as Beethoven, Mozart, etc. Clynes' pulses are not only emotional categories, but also, and perhaps more significantly, gestural utterances. The curves associated with pulses, as shown in figure 13.7, are gestural shapes, movements in the space of duration and loudness. Clynes accordingly constructed and patented a machine, the *sentograph*, providing us with an interface to grasp such gestural movements. Following Clynes' ideas, Hungarian composer Tamas Ungvary has constructed a sentograph that can be used by improvising composers in order to play/create music by gestural input [141]. Ungvary replaces the usual encoding of sound events in discrete points in a parameter space by an intrinsically gestural input that is given by variable pressure and angle on a joystick (figure 14.3). Despite the fascinating perspective on musical creation, the gestural input remains very abstract insofar as no significant movement of the fingers is possible. The musician has to stay in contact with that fixed piece of metal and cannot move freely in space. This restriction heavily limits the natural human need for movements

Fig. 14.3. Tamas Ungvary playing the sentograph. The joystick is accessed with the right middle finger.

when gestures have to be created from the living body. Perhaps a more natural encoding of the input parameters would improve the expressive power of this interesting machine.

14.6 Johan Sundberg, Neil P. McAgnus Todd: Mechanical Models of Gestures in Music

On a more down-to-earth level, gesture has been studied by Johan Sundberg and collaborators. In a paper entitled "Is the Musical Ritard an Allusion to Physical Motion" [64], Sunberg and Ulf Kronman have studied final ritard as a phenomenon akin to physical ritard. The model conjectures that a tempo decrease at the end of a musical piece would be related to a quadratic function, which appears for mechanical ritard with a constant force. So we are given a constant force F, and its action on a given mass m, which generates a constant deceleration $a = F/m$ according to Newton's second law. Given an initial velocity v, the velocity after t seconds is $v - a.t$. Whence the distance $s(t)$ traveled after t seconds is $s(t) = \int_0^t v - a.\tau d\tau = t.v - a/2.t^2$. If the final velocity at time t_0 is 0, we have $t_0.a = v$, whence $s(t_0) = (v/a).v - a/2.(v/a)^2 = v^2/2a$. Therefore velocity at time t is $v(t) = v.\sqrt{1 - s(t)/s(t_0)}$. Supposing that this physical situation relates to the musical one by a constant c, i.e. $s(t) = c.E(t)$, E being the symbolic onset, we get $T(t) = T(t_0).\sqrt{1 - E(t)/E(t_0)}$. In other words,

$$T(E) = T(E_0).\sqrt{1 - E/E_0}$$

namely the tempo at onset E being the above function of the tempo $T(E_0)$ at the beginning E_0 of the ritard, the onset E and the beginning onset E_0. The experimental situation is shown in figure 14.4. The parabolic tempo curve relates to the phase I in the left graphic. Phase II is interpreted as a linear tempo decrease.

Besides the poor fit of the measured tempo with the mathematical curve, the question arises why such a mechanical function should hold. What is the musical analog to mass, what is the force analog to a constant mechanical force? We do not see any musical structure entailing such a mechanical model. It is interesting that the ritard phase II relates to a quite sophisticated harmonic and melodic musical process, which is not taken into account.

Another mechanical model of agogics has been proposed by Neil P. McAgnus Todd in [137]. He rightly observes that the final retard is only a very special agogical situation and therefore models his tempo curves according to a superposition of accelerando/ritardando units that are defined by a triangular sink potential V. Accordingly, tempo is defined as a velocity v, and the total energy of the system $E = \frac{1}{2}mv^2 + V$—supposed to be constant (why so?)—gives the velocity formula $v = \sqrt{2(E - V)/m}$. Todd further supposes that there is an intensity variable I for loudness, with a relation $I = K.v^2$ that is common to many physical systems. This yields the relation $I = 2K(E - V)/m$ and sums up to an aggregated formula $I = \sum_l 2K(E - V_l)/m_l$ if the grouping

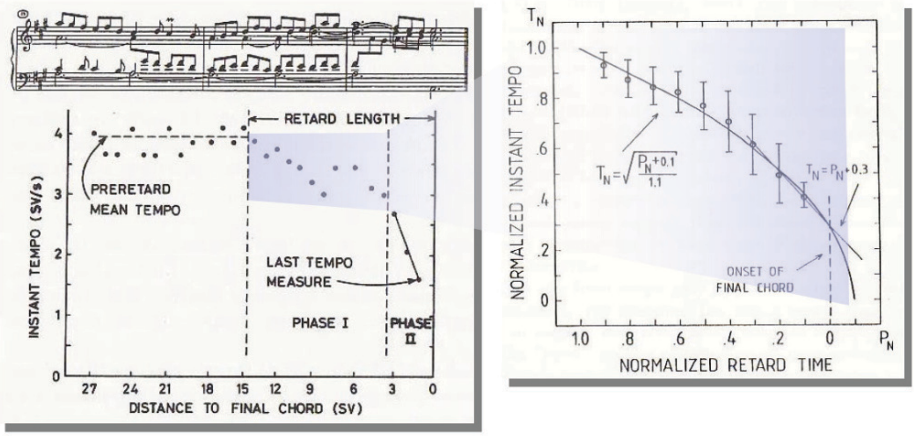

Fig. 14.4. The parabolic tempo curve (right figure) relates to the phase I in the left graphic. Phase II is interpreted as a linear tempo decrease.

of the piece is taken into account. The idea is that there is a physical energy and intensity parameter system that controls the "surface" of the tempo (= velocity) via classical energy formulas. The background structure is an energetic one, i.e., the tempo curve and loudness are expressions of mechanical dynamics. The author comments on his method as follows [137, p.3549]:

> *The model of musical dynamics presented in this paper was based on two basic principles. First, that musical expression has its origins in simple motor actions and that the performance and perception of tempo/musical dynamics is based on an internal sense of motion. Second, that this internal movement is organized in a hierarchical manner corresponding to how the grouping of phrase structure is organized in the performer's memory.*

The author also suggests a physiological correlate of this model (loc. cit.):

> *...it may be the case that expressive sounds can induce a percept of self-motion in the listener and that the internal sense of motion referred to above may have its origin in the central vestibular system. Thus, according to this theory, the reason why expression based on the equation of elementary mechanics sounds natural is that the vestibular system evolved to deal with precisely these kinds of motions.*

Todd refers to the insights of neurophysiologists that the vestibular system is also sensitive to vibrational phenomena. The musical expressivity is therefore understood as an effect of transformed neurophysiological motion.

The drawback of this approach is that finer musical structures are not involved in the structuring of the energy that shapes tempo/intensity. And even if that could be done, there is an essential kernel of this shaping method that

should be based upon paradigms of motion. These paradigms do not however appear clearly in the above approach. More precisely: The complex motion dynamics of the vestibular system cannot easily be mapped onto the structures of performative expressivity. What is the operator that transforms whatever structures of motion into expression parameters? If music were isomorphic to motion, no such isomorphism could be recognized from Todd's approach.

Fig. 14.5. The gesturality of four famous pianists: Vladimir Horowitz (top left), Glenn Gould (top right), Arturo Benedetti Michelangeli (bottom left), Cecil Taylor (bottom right)

14.7 Guerino Mazzola and Stefan Müller: Modeling the Pianist's Hand

One of the most evident gestural expressions is the body movement of a performing musician. We instantly recognize the gestural power of four famous pianists as shown in figure 14.5. It is logical that one should therefore attempt to model gestures of musicians. In collaboration with my PhD student Stefan Müller, we embarked in the modeling of the pianist's hand [85] on the level of computer graphics. The idea was not only to model the hand's movements,

but also to implement a software that could transform the abstract symbols of a score into hand movements that were adequate for the rendering of the score on a piano keyboard.

The project had three components:

1. Modeling the hand with its spatio-temporal trajectory in the movement.
2. Transforming abstract score symbols of notes (what we call deep-frozen gestures, since they historically stem from neumatic abstraction, neumes being gestural signs) into *symbolic gestures*, i.e. curves in a space related to the piano keyboard geometry.
3. Deforming symbolic hand gestures into physically valid spatio-temporal curves of the pianist's hand.

14.7.1 Modeling the Hand

This task was accomplished with a simplified representation of the hand by six curves $\gamma_i(t)$ in physical space-time with space axes x, y, z to denote the momentous position of the hand, and e, the physical time of that position. The curve parameter t is not the physical time, it is just an abstract curve parameter (figure 14.6). The curves $\gamma_1, \ldots \gamma_5$ represent thumb, index, middle, ring, and little finger, respectively, while γ_6 represents the carpus.

These curves are subjected to geometric constraints G resulting from their connectivity as parts of the hand's geometry. We refer to [85] for more details. And the curves are subjected to mechanical constraints M, which means that if we think of the ith finger's mass m_i, and if the pianist is capable of exerting a maximal force of K_i upon that mass, then Newton's second law imposes the inequality

$$m_i |d^2 \gamma_i^{space} / de^2(t)| \leq K_i$$

at any curve parameter value t, where γ^{space} is the three-dimensional spatial part of the curve.

14.7.2 Transforming Abstract Note Symbols into Symbolic Gestures

Refer to figure 14.7 for the following discussion. In traditional performance theory, we look at the transformation $\wp_{score}$ of score symbols into sound events. This is shown in the bottom row of the rectangular diagram of figure 14.7. In the gestural extension of this disembodied process, we have to create the sonic result via gestural actions. The sounds are just the result of physical gesture curves interacting with the keyboard; these curves are shown in the right top corner of the diagram.

In order to generate these physical curves, one first has to unfreeze the note symbols and to transform them into gestural symbols. This unfreezing process is shown in the left top-bottom half of the diagram. This does not create physical gesture curves, but only symbolic curves, which are faithful

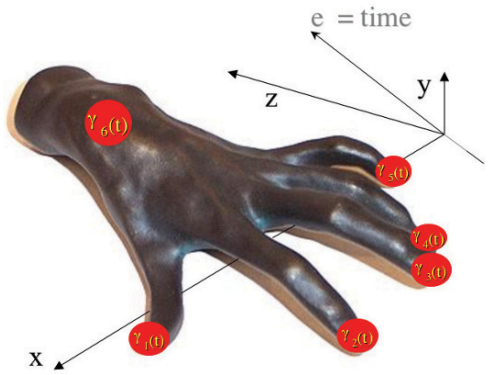

Fig. 14.6. Modeling the pianist's hand.

representations of the note symbols. This process resembles the MIDI interpretation of notes insofar as the commands associated with notes are abstract movements: In MIDI, a note is defined by an ON command, which means go down to that pitch at a determined moment, and the MIDI velocity used to move down to the key defines loudness. Then the finger remains in that position until the OFF command tells it to move up again, etc.

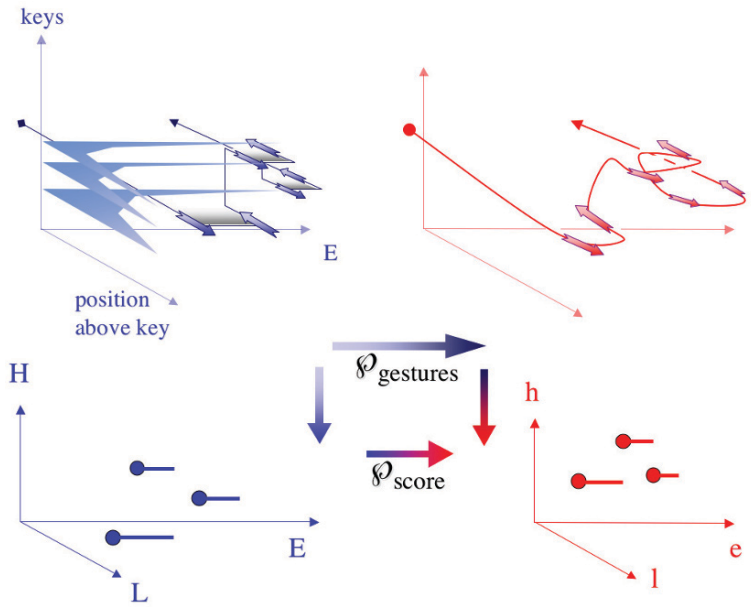

Fig. 14.7. The four levels of performance: symbolic score representation (left bottom), performed sound events (right bottom), symbolic gesture curves (left top), physical gestures (right top).

This representation defines a very abstract curve, but it is this that tells the fingers in a qualitative way how to move. This movement is shown in the left top corner. We see the symbolic gesture associated with a sequence of three notes in the left bottom corner. The finger moves down over a first key, then remains there and after the duration moves up, changes the key coordinate, goes down to the second key with a second velocity, remains there for its duration, moves up, shifts to a third key position, moves down with the third velocity, remains for that duration, and finally moves up. All these phases are connected in a continuous curve, which has angles, i.e., it is not differentiable, and whose movement is orthogonal to the time axis E when moving down at a determined velocity.

14.7.3 Deforming Symbolic Hand Gestures into Physically Valid Curves

The third step toward gestural performance is the horizontal transformation on top of the diagram in figure 14.7. The given symbolic curve does the right thing, but it does not move within the geometric and physical constraints. These constraints define a subspace of the space of all continuous curves, in fact a manifold $X(G, M)$ in terms of global geometry. We are given the symbolic gesture curve from the left top data and now have to create a pysically valid deformation thereof, i.e. one that fulfills the geometric and mechanical constraints G, M. This is a very delicate operation. Essentially, it boils down to looking at the symbolic curve $\gamma_{\text{Symbolic}}(t)$ and then searching for one $\gamma_{\text{Physical}}(t)$ that is as near as possible to $\gamma_{\text{Symbolic}}(t)$ and lives in $X(G, M)$. The delicate point is that it is often not possible to cope with all conditions for a perfect performance, since, for example, physics does not allow for infinite velocities. So when the finger has to play two different keys in immediate succession without a pause, the duration of the first note must be shortened in order to jump from the first to the second key. Such difficulties must be met by defining distance between curves in such a way that musical constraints are given a high weight. For example, the prescribed key coordinates cannot be changed, while durations may be changed, but only a little, etc. It may then happen that there is no solution to a given score input and its associated symbolic curve. This must be possible as a function of the anatomic and physical constraints given from the human conditions. We have implemented this process and have performed a simple Czerny exercise (the same as was used for the Espresso Rubette shown in figure 10.4 in section 10.3), illustrated in figure 14.8.

14.8 A Mathematical Gesture Theory

The desire of a precise definition of a gesture was already in the air following all those more or less philosophical and aethetical approaches to gestures, which we have discussed previously. Following the experimental implementation of a

Fig. 14.8. A Czerny exercise played by the computer-graphical model of a pianist's hand according to the gestural "unfreezing" process described in the text.

performing pianist's hand in [85], a more explicit definition of a gesture became feasible. We also discovered that there was a very interesting definition around that came from medieval studies. Medieval theologist Hugues de Saint Victor defines: "Gestus est motus et figuratio membrorum corporis, ad omnem agendi et habendi modum." Gesture is the movement and figuration of the body's limbs with an aim, but also according to the measure and modality proper to the achievement of all action and attitude. Most important is that it is an articulated figuration, a composition of parts (limbs), and that it includes a movement of that figuration in the space-time of the given body. Moreover, it serves for any (omnem) mode of action and attitude, so it has a purpose or target, but it does not, automatically, point to a semantic level—it only reaches the mode of an activity/habit. So this concept follows Adorno's and even more strongly Wieland's gesture philosophy, which is a spatial one, without including semantics automatically; it is only "expression of expression."

Following this approach, we define a gesture as being (a) a directed graph $\mathcal{D}$, called the gesture's *skeleton*. This is the schematic description of the configuration of limbs in Saint Victor's definition. Then we need (b) a map g that associates with each arrow a of $\mathcal{D}$ a continuous curve $g(a) : I \to X$ defined on the unit interval $I = [0, 1]$ of the real number line in such a way that matching arrows carry over to matching continuous curves. The system of these continuous curves is called the gesture's *body* (figure 14.9). It is in this latter part of the gesture that movement of the gestural configuration takes place—for example, if the space X is a space-time, i.e. contains time as a coordinate, such as we had defined gestural curves in the modeling of the pianist's hand in section 14.7.1. The space of all continuous curves in X is denoted by $\overrightarrow{X}$, meaning that it is a (big) directed graph, whose arrows are the continuous curves, and whose vertexes are the points of X. Then the gesture can then be written as a map of directed graphs $g : \mathcal{D} \to \overrightarrow{X}$.

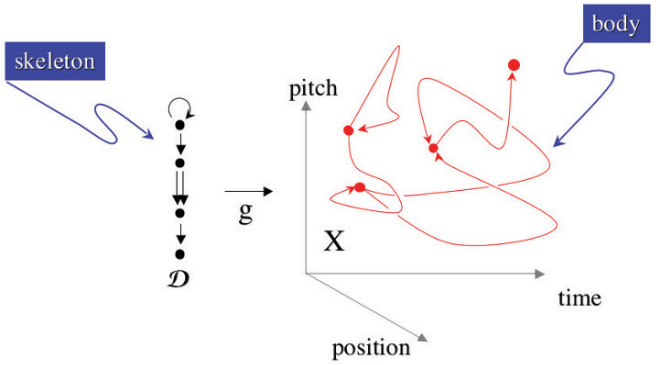

Fig. 14.9. A gesture is a map g from the digraph $\mathcal{D}$ (the skeleton) into the system of continuous curves in a topological space X, the gesture's body. In this example, X is the space with coordinates used for the fingertip positions of a pianist's hand at a given key (pitch), a level above the key (position), and the time of this event.

Two gestures $g : \mathcal{D} \to \overrightarrow{X}$ and $h : \mathcal{E} \to \overrightarrow{Y}$ can be related to each other by a kind of function called *morphism of gestures*, which is a pair $u : \mathcal{D} \to \mathcal{E}, v : X \to Y$ with u being a map of directed graphs, and v being continuous, such that $h \circ u = \overrightarrow{v} \circ g$; see figure 14.10 for an example of such a morphism that connects gestures in multitouch spaces to gestures in musical spaces. This situation is typical in gesture theory: We have gestures in different topological spaces and need to connect them via auxiliary maps between their skeleta and/or between the topological "carrier" spaces.

The most powerful device in this mathematical theory of gestures is the concept of a hypergesture. Recall from section 14.4 that Renate Wieland mentioned that mysterious idea of a "the gesture within the gesture." This construction is now fairly easy: It can be shown that the set of gestures from a fixed skeleton $\mathcal{D}$ to a fixed topological space X is itself a topological space, denoted by $\mathcal{D}\overrightarrow{@}X$. Therefore we may consider gestures $h : \mathcal{F} \to \mathcal{D}\overrightarrow{@}X$. Such gestures are called *hypergestures*: gestures whose body is a system of curves of gestures! Although this sounds complicated, it is quite intuitive. For example, if we have a gesture with a loop as a skeleton, and then a hypergesture with again the loop as skeleton. Then the hypergesture is a loop of loops, in fact a nice closed tube, as in figure 14.11.

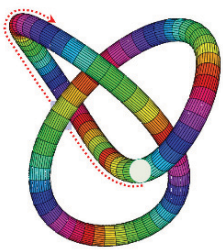

Fig. 14.11. A closed tube is a hypergesture, namely a loop of loops.

Such hypergestures are very useful in generating gestural interpretations of classical musical compositions. For example, Beethoven's tonal modulation $B_\flat major \to Gmajor$ in the beginning of op.106, Allegro, can be described by use of such hypergestures, see [93].

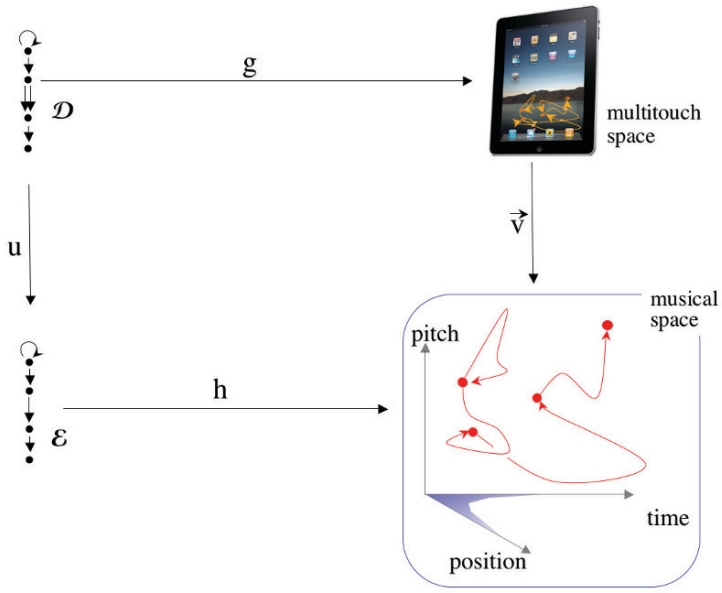

Fig. 14.10. Gestures g and h may also be connected to each other by a so-called *gesture morphism*, which means that the digraphs and the topological spaces are provided with morphisms, respectively, that are compatible with the two gestures (commutative diagrams).

14.9 Anders Friberg, Antoni Camurri et al.: Computer-aided Emotional Analysis of Performance

We terminate our tour d'horizon of gestural expressivity with a short review of recent attempts to incorporate gesture and emotion into the analysis of performance.

In research related to a EU program (MEGA) [36], [12], Anders Friberg, Antonio Camurri and collaborators implement a software that analyzes emotions from music and body gestures and then generates an emotional valuation thereof, yielding a fuzzy mixture of happiness, sadness, and anger. The system takes cues from music tempo, dynamics, and articulation, from body motion, and from overall motion (via video displacement pixels), and valuates the motion and music performance cues (large, fast, uneven, jerky for motion/anger, etc.; loud, fast, staccato, sharp timbre for musical anger, etc.). Then after a gauging/standardization step, the three fuzzy values (0 to 1) of the three emotions are calculated (by the fuzzy mapper, discussed in section 13.5, see also figure 13.10). The expression mapper can be used to visualize the emotions by color and shape (see figure 14.15 for its process flow chart). This visualization was implemented by Camurri and is shown in figure 14.12. The gestural processing of Camurri's *video mapper* takes motion cues from a pixel image as

shown in figure 14.13. Three motion cues are determined: First it counts all visible pixels and determines the bounding rectangle defining the area of the picture that contains all non-zero pixels. The instant width and height of the bounding rectangle are computed and their peak-to-peak amplitude variations constitute the cues width-p and height-p. Second, the maximum velocity of gestures in the horizontal plane is calculated, followed by a calculation of the time between gestures in the horizontal plane—the third one.

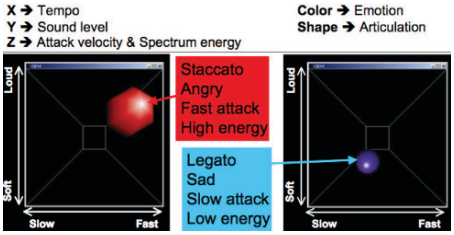

Fig. 14.12. Antonio Camurri's expressive mapper visualization coined *expressiball*.

Fig. 14.13. The video mapper input takes essentially the rectangle around the visible pixels.

Players used the system in the 2004 group game Ghost in the Cave at the Tekniska museet Stockholm (figure 14.14). The game is played in two teams, each with a main player. Each team's task is to control a fish avatar in an underwater environment, moving it into three different caves. In the caves are ghosts expressing different emotions. The main players have to express the same emotion, causing their fish to change accordingly (figure 14.14).

Fig. 14.14. The Ghost in the Cave game at the Tekniska museet Stockholm.

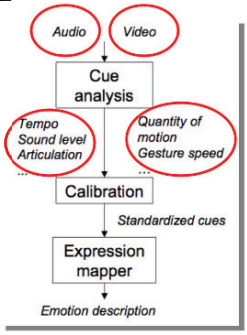

Fig. 14.15. The audio-video processing flow chart.