Sculpting a Sound Space with Information Properties

Stephen Barrass
stephen.barrass@dit.csiro.au
CSIRO Division of Information Technology,
GPO Box 664, Canberra,
ACT 2601, Australia

Prepress Draft.

Citation:
Stephen Barrass. 1996. Sculpting a sound space with information properties. Org. Sound 1, 2 (August 1996), 125-136. DOI=http://dx.doi.org/10.1017/S1355771896000040

ABSTRACT
This is a description of the design and construction of an organised sound space to support information representations in the human computer interface. The design of the sound space is guided by 4 principles which match perceptual structure with data structure to improve natural comprehension of an auditory display. These principles - completeness, comprehensibility, consistency, and cohesiveness have been generalised from the use of colour displays in scientific visualisation. The choice of perceptual parameters to represent different types of data is informed by the body of psychoacoustic literature.
The raw material for the construction of the sound space is the McGill University Master Samples (MUMS) palette of musical instrument samples. This is an important choice because this reference resource enables reproduction and confirmation of the results. The construction was carried out in four stages - the "pedestal", the "skin", the "skeleton" and the" flesh". The pedestal consists of 8 equally discriminable timbres organised in a circle by perceptual similarity. The skin is the boundary of variation in the space, defining the limits of dynamic range for pitch and brightness at each timbre. The skeleton characterises the internal behaviour of the space at a number of perceptually measured points. The flesh is a continuous medium moulded to the skeleton and skin, realised by a 3D regularised linear spline interpolation.
The concrete realisation of the sound space can be investigated through a user interface, called the GamutExplorer. Colour visualisations of slices and wireframe views of the 3D space can be chosen, and sounds can be picked with a mouse.

Keywords: Sonification, auditory display, psychoacoustics, visualisation, human-computer interface, organised sound

1. INTRODUCTION

Sounds are a form of information about the world. Sounds tell us about the sizes of things, what they are made of, where they are, and what they are doing. We can compare distances, speeds, or densities by listening. Patterns of sounds allow us to monitor patterns of activity at a remote location. The informational properties of sounds help us to carry out our everyday activities. Since many activities now also involve computers, perhaps there is a potential to use sounds to help with these types of activities too? In order to use sounds in this new way we must first understand how people hear information in sounds, and how different aspects of

sound carry different types of information. This will enable the design of useful auditory displays which can support people involved in computer-based activities.
There are a number of auditory display techniques currently in use, which are described by Kramer (Kramer 1992) who has arranged them along a spectrum from analogic to symbolic representation. At the analogic end is the audification method in which the data are directly converted to sound samples. This method is particularly useful for seismic and other data types which are closely related in form to sound waves. Toward the symbolic end are auditory icons and earcons - two well known methods which have been applied to enhance graphical user interfaces by signifying events such as files being opened or mail arriving. Auditory displays based on these techniques have demonstrated that sounds can provide extra information in an interface. However it is not always obvious which technique to use for a particular information display, because none of these techniques explicitly defines what type of information it can carry.

In this paper a new method of auditory display which directly addresses the issues of data types and their perception through sounds is proposed. This data-sensitive perceptually-based methodology is grounded in 4 principles of perceptual display which have been abstracted from the use of colour displays in scientific visualisation. These principles are used to design the organisation of a sound space to support a general range of information representations. Psychoacoustic research results and models were used to inform the assignment of perceptual dimensions so that the perceptual structure is aligned with the information structure of the space. This method has the capability to span a broad range of representations, from near analogic to symbolic.

In order to verify the design a concrete instance of the sound space was constructed. The McGill Master Samples (MUMS) palette of musical instrument samples was specially selected as the raw material for construction (Opolko and Wapnick 1995). This choice is important. The MUMS samples are of high quality and were intended for perceptual research, they are readily available (http://lecaine.music.mcgill.ca/MUMS/Html/Mums.html), and their value is enhanced by spectral analysis data which is also readily available (http://www.parmly.luc.edu/sandell/sharc/SHARC.homepage.html). These samples provide a reference point for other researchers who may repeat or confirm the results of experiments which employ these complex time varying sounds.

The next section describes the principles of organisation of the sound space. The design of the space based on psychoacoustic literature is then detailed. The final section is an account of the realisation of a concrete instance of the sound space.

2. ORGANISATION

The representation of information in computer displays has been an important problem in the field of scientific visualisation. The use of colour in these representations has drawn on research in colour perception, and a methodology of colour display based on the alignment of data structures with perceptual structures has been developed (Robertson 1988, Rogowitz 1993).

The principles which underly this method of perceptual information display are:

- completeness - it is able to represent a general range of data types
- comprehensibility - it makes it possible to comprehend the structure of the data
- consistency - representations change in a manner consistent with changes in the data

· cohesiveness - a representation forms a cohesive perceptual entity

2.1. COMPLETENESS

To be useful for a general range of tasks the information display must be capable of representing a complete range of data types. Data types can be classified according to the following scheme (Weirsma 1991):

· nominal - categories e.g. weather types = summer, spring, autumn, winter
· ordinal - categories with order e.g. weather forecast = fine, fair, foul
· interval - continuous order and a unit- e.g. temperature = 10, 20, 30 degrees celsius
· ratio - continuous order, a unit and a natural zero e.g. rainfall = 0, 10, 20 mm

Note that a "natural zero" reflects a total absence of whatever is being measured, and all scales start from this point independent of their unit. For example a reading of zero rainfall has the same meaning whether the scale is mm or inches.

2.2. COMPREHENSIBILITY

A comprehensible display enables a person to understand the structure of the data it represents (Ware 1993). The construction of a such a display requires the alignment of perceptual structure with data structure. The body of psychoacoustic literature provides insights into different aspects of auditory perception which can inform this alignment. Of particular interest is the classification of sensations as either continuous or categorical. Continuous sensations have an observable unidimensional order, meaning that the level of the sensation can be judged and compared in a quantitative way. Some examples are sound loudness and colour lightness. Categorical sensations have difference but no observable order. These compound sensations have many aspects of variation which underly the ability to discriminate, segregate and partition sensory inputs. Some examples are sound timbre, visual texture and haptic shape. Comparisons can be made in terms of similarity or difference in some subset of dimensions, in a qualitative manner. The continuous/categorical distinction corresponds well with the data types described in the previous section. This leads to the proposal of a representation scheme in which continuous perceptions are used to represent continuous data types and categorical perceptions are used to represent categorical data types. For example, the categorical perception of timbre is suited to the representation of nominal data, so that spurious inferences of order are not heard in the display.

The link between data structure and perceptual structure has been facilitated by extending the perceptual types as follows:

· categorical - has difference but no order e.g. timbre - flute, violin, cello
· categorical with order - has difference and order e.g. pitch classes - a,b,c,..
· continuous - has order and a unit of equal difference e.g. diatonic pitch scale - 1,2,3 semitones
· continuous with natural zero- has order, unit and a natural zero e.g. loudness - 0,10,20 sones

This extension allows the linking of data types and perceptual types as shown in Table 1.

Data Type	**Perception Type**

Nominal	Categorical
Categorical with Order	Ordinal
Continuous	Interval
Continuous with Zero	Ratio

Table 1: Aligning Data Structure with Perceptual Structure

2.3. CONSISTENCY

In order to preserve the interrelationships between data values mapped into sounds it is necessary that there be a consistent correspondence between data variation and perceptual variation throughout the display space. For example, an equal change in pitch corresponds with an equal change in the data, independent of the actual pitch range used.

Psychophysical scales provide a means of making consistent data to sound mappings for continuous data types. Various scales have been constructed using techniques such as fractionation, just noticeable differences, and magnitude estimation. Some examples are the Sone scale of loudness, the Mel scale of pitch, and the Acum scale of sound brightness (Zwicker 1990). These scales have a unidimensional order and a unit of equal perceptual change.

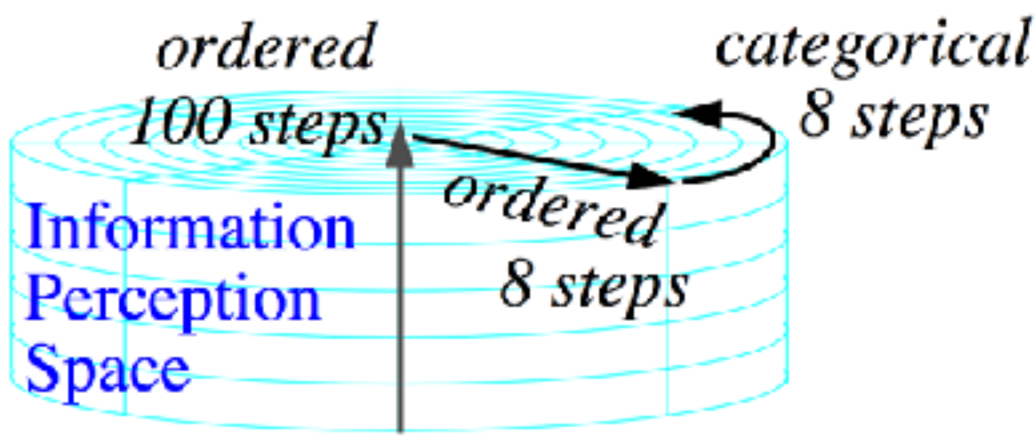
The consistent representation of categorical data relationships requires the ability to perceive equally discriminable categories of sounds. Uniformly scaled euclidean spaces which describe the relationships between categorical perceptions have been constructed using the multidimensional scaling technique (MDS). Uniformly spaced points can be selected from these spaces to form perceptual categories with defined relationships.

2.4. COHESIVENESS

Perceptually cohesive sound complexes are groups of sounds which are heard as a unified entity that can be selected for individual attention from a background of other sounds in an auditory scene. The control of cohesiveness supports the simultaneous display of multiple sound complexes, through control of perceptual interactions between auditory elements.
The grouping and segregation of sounds in a concurrent auditory array has been the subject of experiments and theories of auditory scene analysis (Bregman 1990). Perceptual interactions may be controlled by using auditory streaming effects to bind groups of sounds together into cohesive complexes. To achieve this control it is necessary to choose perceptual dimensions which have strong grouping effects; for example streaming experiments have shown that timbre, pitch, spectral centroid, synchronicity of partials, presentation rate and many other factors have various degrees of influence in grouping.

3. DESIGN

Having established the principles of completeness, comprehensibility, consistency, and cohesiveness, we must somehow harness them to the problem of creating auditory displays. Rules and guidelines can be helpful in solving problems, but a key aspect of user interfaces is the leverage of knowledge through cognitive artifacts which alter the focus of the task in a way which enhances the skills of the designer. An example is the Hue, Saturation, Lightness (HSL) colour space which allows navigation and selection of colours which are organised in a geometric 3D arrangement by perceived similarity in naturally ordered dimensions of perceptual variation (Barnard 1991, Norman 1991). This colour space has had an important role in scientific visualisations in which data structure is displayed through colour representations. The HSL colour space supports the principles of completeness, comprehensibility and consistency in a manner which is easy to use and interact with. Completeness is supported because hue perception is categorical, while saturation and lightness are continuous. Comprehensible colour mappings have been defined using the space by matching data structure with the structure of colour perception (Robertson 1988, Rogowitz 1993). Consistency is supported by uniform scaling of the colour space (Robertson 1988). By considering the structure of the HSL space we can induce a template for representing data in perceptual displays. The Information-Perception Space that can be abstracted in this way is shown in Figure 1. This cylindrical polar space has a categorical angular dimension, a continuous radial dimension and a continuous vertical axis dimension. Other arrangements could be used but the structure of the HSL space provides us with a familiar and well-known starting point.



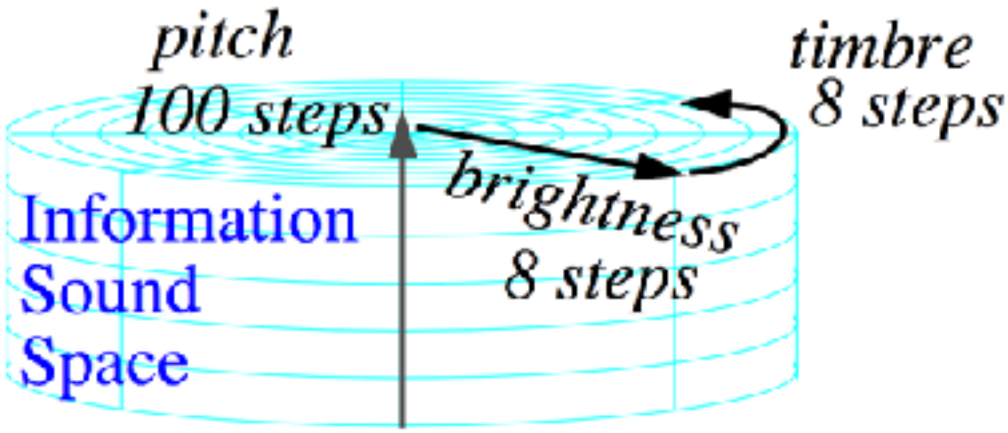
[Figure 1. Blueprint for an Information-Perception Space]

3.1. INFORMATION-SOUND SPACE

An Information-Sound Space can be derived from the abstract Information-Perception space by assigning aspects of sound perception to the perceptual axes in accordance with the organisational principles. An example of one possible derivation, called the TBP Sound Space, is shown in Figure 2, but there are many other permutations which could also satisfy the criteria. The TBP Sound Space has a categorical timbre angle (T), continuous brightness radius (B), and continuous pitch axis (P). Note that the motivation for the space is not to describe hearing perception (as was the case in the development of the HSL colour model), but to support a method of data-sensitive auditory display.

Timbre has been chosen for the categorical angle because of its categorical nature (Ehresman and Wessel 1978), salience in perceptual grouping (Bregman 1991), and availability of scaled results from a number of MDS studies. In the case of the continuous axes the contenders need to be independently observable across the range of timbres, to have a continuous scale, and to have salience in perceptual grouping. The main candidates are loudness (sones), pitch (semitones) and brightness (acums (Zwicker and Fastl 1990)). The selection of pitch and brightness was made because they have strong, independent and additive effects in perceptual grouping, whereas loudness has only minor effects (Bregman 1991) and does not support cohesiveness well. Brightness is strongly related to timbre but the independence of brightness is supported by the observation that sound engineers commonly apply EQ filters to instrument sounds in recording studios without altering the categorical identity of the instrument. Most studies of timbre agree on 2 major orthogonal aspects of timbre - the attack segment of the temporal envelope, and the centroid of the spectral energy distribution (Grey 1975, Plomp 1976). The term "timbre category" will be used in this paper to refer to the perception of timbre difference due to the attack segment of a sound. The term "brightness" will be used to refer to the continuous component of timbre related to the spectral centroid.

At this point it should be noted that none of these aspects of sound are truly orthogonal. Even pitch and loudness, which are considered the most independent aspects, have influence on each other - so for example a change in pitch can affect perceived loudness even though the sound intensity remains constant (Zwicker and Fastl 1990). Loudness will vary considerably throughout the space, and this may be corrected using loudness calculation algorithms such as that described in Zwicker and Fastl as ISO standard 532B (Zwicker and Fastl 1990).

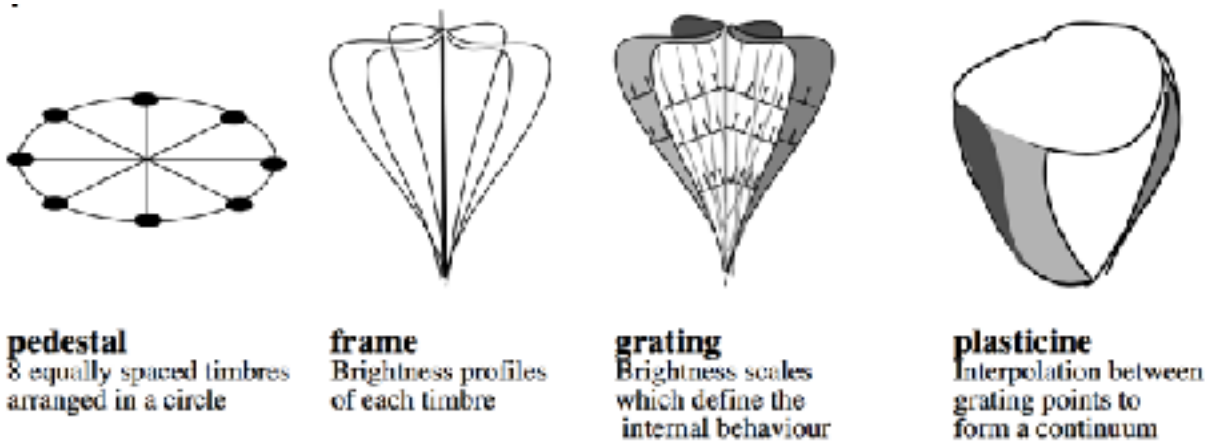


[Figure 2. Blueprint for the TBP Information-Sound Space]

4. CONSTRUCTION

This section uses the design of the TBP Sound Space as a blueprint for the concrete realisation of a general purpose Sound Space with Informational Properties. The raw material of construction is the McGill University Master Samples (MUMS) compact discs of musical instrument samples. These high quality 44100 kHz stereo samples were intended for auditory research work with timbre. There is a different sample for every pitch in the playing range of each instrument, and all of the instruments of the classical orchestra are included. In the past auditory researchers have had to employ simple tones and noises because they are easy to specify and generate, allowing repetition and verification of results by other researchers using different equipment. It has been argued that these are very special types of sounds which are not common outside the laboratory and generalisations of results obtained with these stimuli are questionable. Complex time varying sounds have only recently become describable and transportable, through digital sampling and playback. At last it is possible to use real world sounds. However no two samples are identical so it is necessary to either make the samples from an experiment available in some way, or to use a reference palette. The MUMS samples are readily available and serve as such a reference resource for this experiment. The value of the MUMS palette is significantly enhanced by the SHARC timbre database (Sandell 1995), which contains spectral analyses of most of the samples.

The Concrete Sound Space was built in four stages, as shown in Figure 3 - a "pedestal" of support, a "skin" which defines the boundary, a "skeleton" which defines the internal behaviour, and the "flesh" which is a continuum moulded to the skeleton and contained by the skin.



[Figure 3. Stages of construction of the Concrete Sound Space]

4.1. THE PEDESTAL

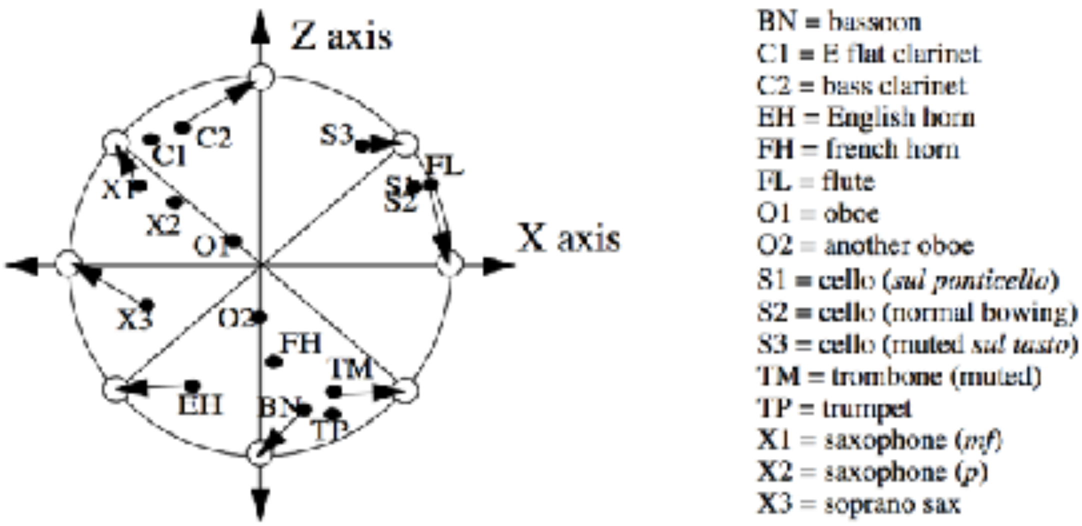
The pedestal is the central organising structure of the space. It is a circle of 8 equally discriminable timbre categories which can represent categorical data types. Different timbres may be substituted into the arrangement if they meet the criteria of equal perceived difference between neighbours. Any group of sounds may be used to build a timbre circle, for example "natural" sounds or "machine" sounds may be selected for semiotic linkage with an application, or several different sound palettes might provide a basis for different timbre circles which are swapped to give modal feedback. In this instance the MUMS musical instrument timbres will be used as the source palette. In order to construct the pedestal it is necessary to measure the similarity between the timbres in some way. As mentioned previously, there have been quite a few MDS studies of timbre which provide perceptually scaled results depicting similarity relationships between particular subsets of timbres. MDS has been criticised because the results are based on small sets of stimuli, the scales of measurement are prescribed and therefore biased, it is difficult to identify the physical correlates of the dimensions of principal variation, and new stimuli cannot be inserted into the results without a global restructuring (Pachella, Somers and Hardzinski 1981). Despite these problems there have been some convincing correspondences between MDS and other methods of timbre analysis in terms of identifying principle components. An alternative would be to use the streaming method (Bregman), which has the advantage of taking into account the simultaneous and sequential aspects of auditory perception and so includes elements which support control of cohesiveness. Unfortunately experiments which measure the effect of the attack segment on grouping have not been done. Bregman conjectures that such an experiment could be carried out by interleaving a familiar tune in one timbre with distractor tones of another timbre, and measuring the rate of presentation at which segregation by pitch destroys the tune as a perceptual unit. Having highlighted this possibility, let us use the available MDS results for the present.

One of the best known MDS studies is Grey's scaling of 16 re-synthesised instruments. The MUMS palette has sampled instruments which correspond with all of those used in Grey's experiment, as shown in Table 2. The validity of this substitution might be questioned since a repetition of the MDS experiment using MUMS samples in place of Grey's resynthesised timbres would likely produce different results. However this does not invalidate a categorical substitution where the criterion is relaxed so that the primary structuring is on equal similarity between neighbours rather than on euclidean distance between all pairwise comparisons. The goal of the exercise is to enable the representation of categorical data relationships using the categorical nature of timbre perception, and the substitution is not of timbres but of timbre categories. The categorical substitution can be justified under the assumption that sounds which originated from similar physical sources (e.g. two different cellos) are more perceptually similar than sounds from sources as physically different as musical instrument families (e.g. a cello and a flute). This remains satisfactory because categorical difference is the essential characteristic required for representing nominal and ordinal data.

Grey’s Timbre	**MUMS sample**
Flute	flute, vibrato
Cello (muted sul tasto)	Cello muted with vibrato
Bass clarinet	Bass clarinet
Saxophone	Tenor saxophone
Sopranos sax	Soprano saxophone
English Horn	English horn
Bassoon	Bassoon
Trombone muted	Tenor trombone, muted

[Table 2: Matching Grey's resynthesised timbres with MUMS samples by physical source]

The results from Grey's MDS study of 16 re-synthesised musical instruments were presented in a series of graphic visualisations which showed where each data point lies in the 3 dimensional perceptual space. These results provide an opportunity to select a set of equally discriminable timbres because they are scaled according to identified perceptual axes. The dimensions of principal variation were analysed in terms of the spectrograms of the data points, and it was found that the Y axis was related to spectral energy distribution, whilst the X and Z axes were related to temporal aspects of timbre, covarying with synchronicity in the development of upper harmonics and the presence of low-energy high frequency noise during the attack segment. The procedure for selection of a timbre circle from this perceptually measured palette is shown in Figure 4. A circle which encloses the projection of the data points in the temporal plane is divided into 8 segments of 45 degrees, and the position of each 45 degree increment around the circumference is nominated to represent the categorical timbre of that segment. Because distance is a measure of similarity, the data point in the segment lying closest to each of the equally spaced points on the circumference is allocated to that point. There are only a limited number of data points available to choose from, so that in the segment where there is no data, the closest point from the adjacent segment was used (i.e. TM). This is only a first approximation to equal spacing as can be seen by the small difference in distance between FL and its neighbour S3 on one side, and the much greater distance to its other neighbour TM. This is a consequence of the sparsity and unevenness of the palette, which might be addressed by the use of a different set of sounds.



[Figure 4. Equally Spaced Timbre Circle constructed from Grey's temporal plane]

4.2. THE SKIN

The skin defines the boundary of available variation within the space. The skin is constructed by welding the brightness profile of each timbre to the central seam of the pedestal to form a hollow irregular ball-like shape. This shape represents the limits of pitch and brightness for each timbre in the pedestal. The raw material for the skin is the brightness profile which describes how brightness varies with pitch. Each musical instrument has a unique brightness profile which is reflects the spectral variations at each pitch due to its physics.

In order to make the skin for the TBP Sound Space the brightness profile of each GreyMUMS timbre in the pedestal must be obtained in some way. Relative brightness may be measured using perceptual streaming, and this method was used in a previous experiment with a different palette of musical instrument samples (Barrass 1994b). However these measurements are very time consuming, and require more formal statistical verification across a survey of individuals to be generalised. This led to the consideration of a computational model, shown in Equation 1 (Zwicker and Fastl 1990), which could be used to calculate the brightness of an energy distribution. This model has the advantage that the results are calibrated in terms of a standard unit, called the acum, which is defined as the brightness of a 1 critical bandwidth noise centred at 1 kHz with a sound pressure level of 60 dB.

The calculation requires spectral data for the brightness profiles of each GreyMUMS timbre. The SHARC timbre database contains this data for most of the MUMS samples. The process of selecting a representative spectrum form a time varying sound is described by in the SHARC documentation (Sandell 1995):

- The sound file was converted from 44100 Hz to 22050 Hz, and analysed with a Phase Vocoder (PV).

- The longest continuous stretch of time in which the note was at 75% or more of its maximum amplitude was identified from the PV information. This located the steady portion of the tone.
- An average spectrum was calculated from all the PV frames identified in previous step. Then least squares was used to find the actual PV frame most closely resembling this average spectrum. The point in time corresponding to this PV frame was designated the "representative point".

The brightness of these representative spectral frames was calculated using Equation 1.

$$Acum = 0.11 \frac{\int_0^{24Bark} N g(z) dz}{\int_0^{24Bark} N dz}$$

(EQ 1)

The upper integral is the first moment of specific loudness over the critical band-rate spectrum, where Bark is the unit of critical bandwidth. The lower integral is the overall loudness of the spectrum. N is the summed energy in each critical band, and g(z) is a weighting factor which accounts for the strong increase in brightness at high frequencies. A graph for g(z) is given by Zwicker, but no equation. In order to make calculations a first order approximation was made by measuring values from Zwicker's graph at 1 Bark intervals in the region from 17 to 24 Bark where the weighting takes effect, as shown in Table 3.:

Bark	0-16	17	18	19	20	21	22	23	24
g(z)	1.0	1.1	1.2	1.3	1.5	1.7	2.0	3.0	4.0

[Table 3: Critical band rate weighting factor g(z)]

The calculation requires the spectral data to be in band-rate form, which models frequency spacings along the basiliar membrane. The linear frequency spectra of the SHARC data can be converted to band-rate, prior to the brightness calculation, using Equation 2 (Zwicker and Fastl 1990). The unit of band-rate is the Bark, frequency is in kHz, the angles returned from the arctan expressions are in radians. When Barks is an integer, f is the frequency of the dividing line between critical bands.

*Barks = 13 * arctan(0.76f) + 3.5 * arctan(f*f/56.25)* (EQ 2)

The brightness profile calculated for each GreyMUMS timbre is shown in Figure 5. Note that the saxophones have dummy profiles set to 1 acum because, as yet, there is no SHARC spectral data for these particular instruments.

[

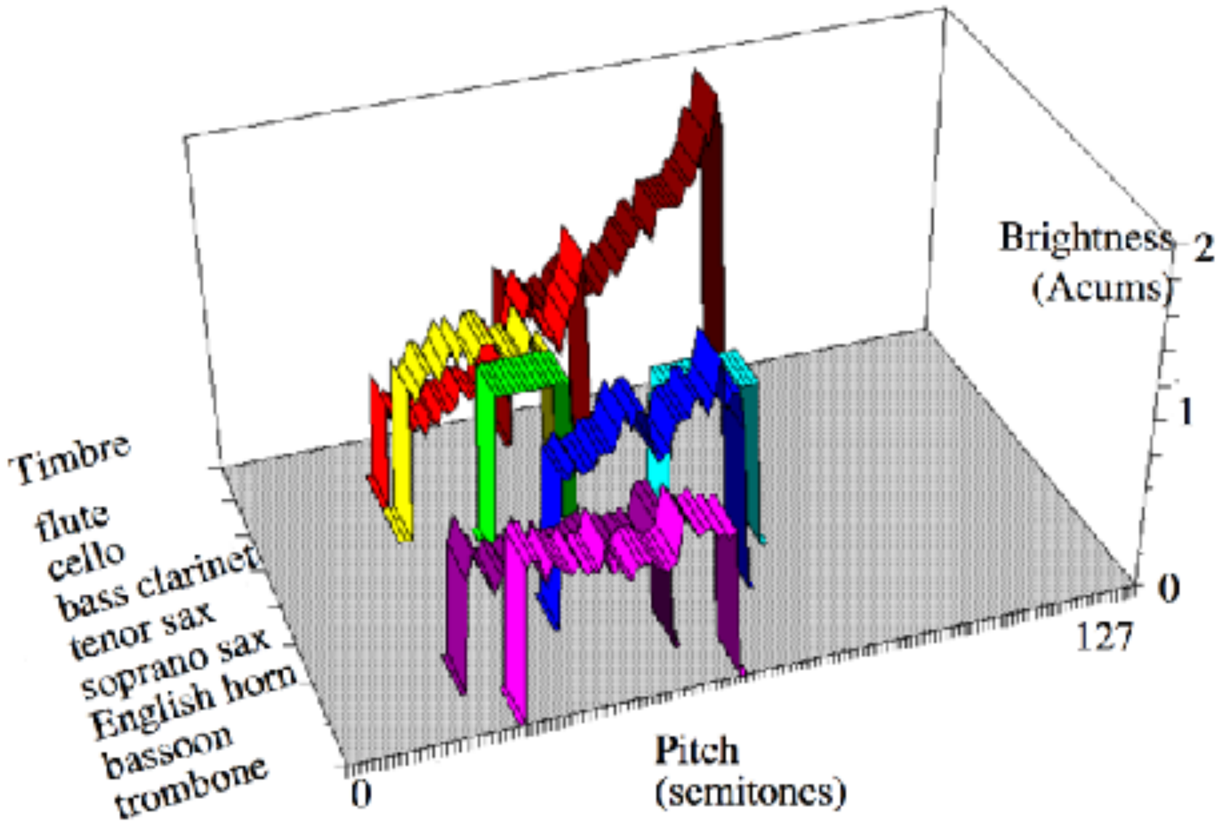


Figure 5. Brightness profiles of GreyMUMS timbres]

4.3. THE SKELETON

The skeleton defines the internal behaviour of the sound space. It consists of a grid of points which define regular steps in brightness, pitch and timbre inside the skin of the Sound Space sculpture. The construction is carried out by measuring a brightness scale for each timbre in the pedestal. A 2D skeletal segment is extrapolated for each timbre by calibrating this scale to the brightness profile over the range of pitches. Each segment has a different shape due to its particular pitch range, brightness profile and brightness scale. The 3D skeleton is completed by arranging the skeletal segments together inside the skin, like pieces of a mandarin. This section will describe the process of measuring the brightness scale and using it to create a skeletal segment.

The psychophysical technique of fractionation can be used to build scales using perceptions of half and double, for example the Mel scale of pitch and the sone scale of loudness. It has also been applied to build an ordered scale of equal brightness steps for various static spectra (Von Bismarck 1974). Through informal investigation it was found that the brightness of the time varying MUMS samples can also be systematically varied by applying a low pass filter to attenuate the upper frequencies. This led to the application of the fractionation technique to create brightness scales for the GreyMUMS samples. All scales were measured at pitch 48, except for the soprano sax which was measured at the beginning of its pitch range at pitch 62. The maximum brightness of each scale differs for each timbre. This maximum brightness was divided into 8 equal steps by fractionation. The dependent parameter had a range of 0..127 which adjusted the cutoff frequency of a first order low pass filter in a linear manner between the fundamental frequency and the Nyquist frequency (22500 kHz) of each sample. The results of the experiment are shown in Figure 6, in which each line represents the relationship between filter cut-off and equal brightness steps for a particular GreyMUMS sample. These results are merely indicative, and were obtained for only one subject (the author). The purpose in showing the data here is to illustrate the construction process. The results were confirmed by a repetition of the experiment a week later.

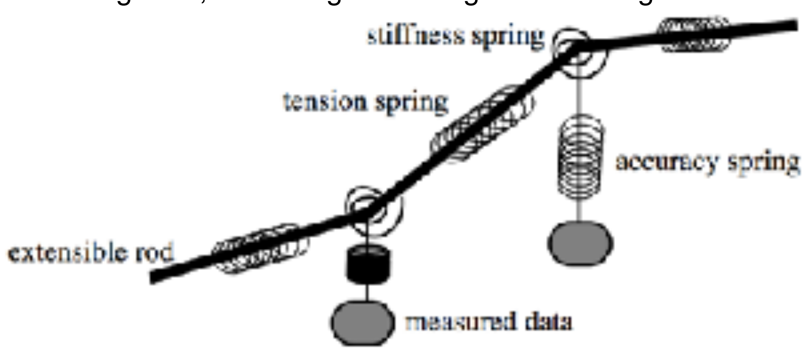
[Figure 6. Filter position versus Brightness step for each GreyMUMS timbre]

The scales are internally consistent but the maximum brightness (level 8) is not calibrated across the scales. This calibration can be carried out using the calculated brightness profiles which correspond with brightness 8, and are in referenced units of acums. The skeletal segments are built by interpolating the calibrated brightness scale over the pitch range of the brightness profile. This results in a 2 dimensional grid of equal brightness and pitch steps at each timbre. These grids are arranged in a 3 dimensional polar layout to form the 3D skeleton of measured data points.

4.4. THE FLESH

The flesh is a continuous `material' moulded to the skeleton which defines the internal behaviour of the space. Its function is to create a continuum of sounds between the measured points.

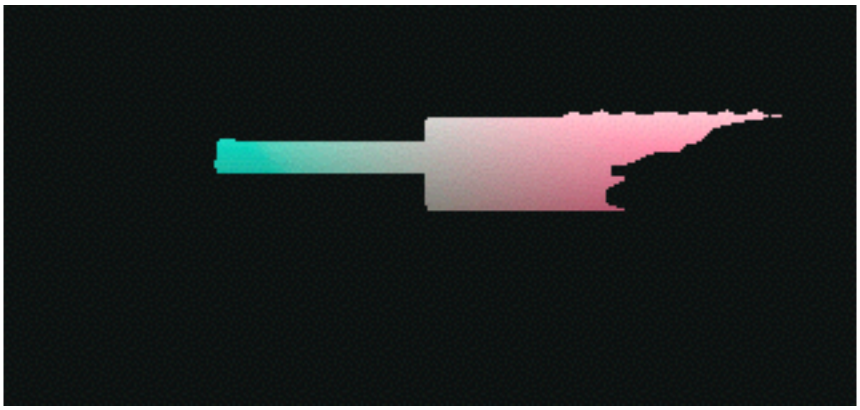
The modelling method was designed to represent the behaviour of a colour output device from discrete measurements of its behaviour in a perceptual colour space (Bone 1993). It consists of a 3D thin plate spline fit to the sparse data points. A 1D mechanical analog of the spline fitting method is shown in Figure 7, and it might be imagined as being a bit like a spring mattress.



[Figure 7. 1D mechanical analog of the spline fitting technique]

This system of extensible rods joined to form a regular mesh has a spring at each intersection node which affects the amount of curvature between them. Each node is constrained to move along a vertical line. The sparse data values are attached to this flexible mesh by another set of springs, and the system minimises a global measure of the curvature of the space. This model of the behaviour can be used to calculate intermediate values between the original samples using a 3D linear interpolation. The technique is effective for fitting continuous non-linear multidimensional data - but the sound space skeleton posed a problem because of the discontinuity between the categorical segments which have to be glued together. The stretchiness of the rods and energy of the springs are parameters which can be adjusted globally at the start of the process to affect the behaviour of the spline material. It was necessary to carry out a systematic search on this parameter space to find a material with properties which could accurately represent the skin and skeletal behaviour of each segment. The size of the margins between timbres was increased so the interpolation technique could flow more smoothly across the them.

The results of the modelling are shown in the following sequence of figures, which depict vertical slices through the sound space together with the calculated skin profile for each segment. It can be seen that there is a good match with the calculated shape.



[Figure 8. Timbre Leaf - soprano sax / flute]

Figure 9. Timbre Leaf - English horn / cello]

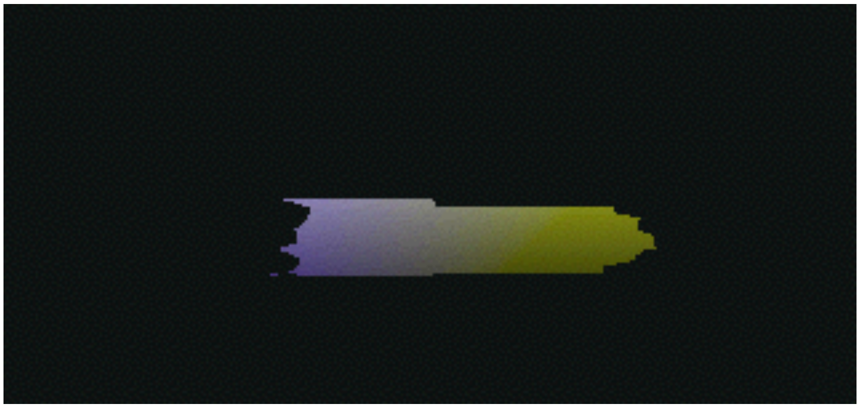
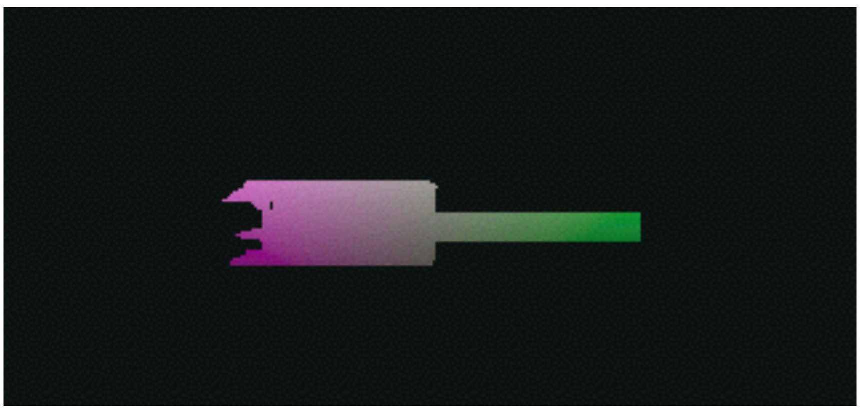


Figure 10. Timbre Leaf - bassoon / bass clarinet]



[Figure 11. Timbre Leaf - trombone / tenor sax]

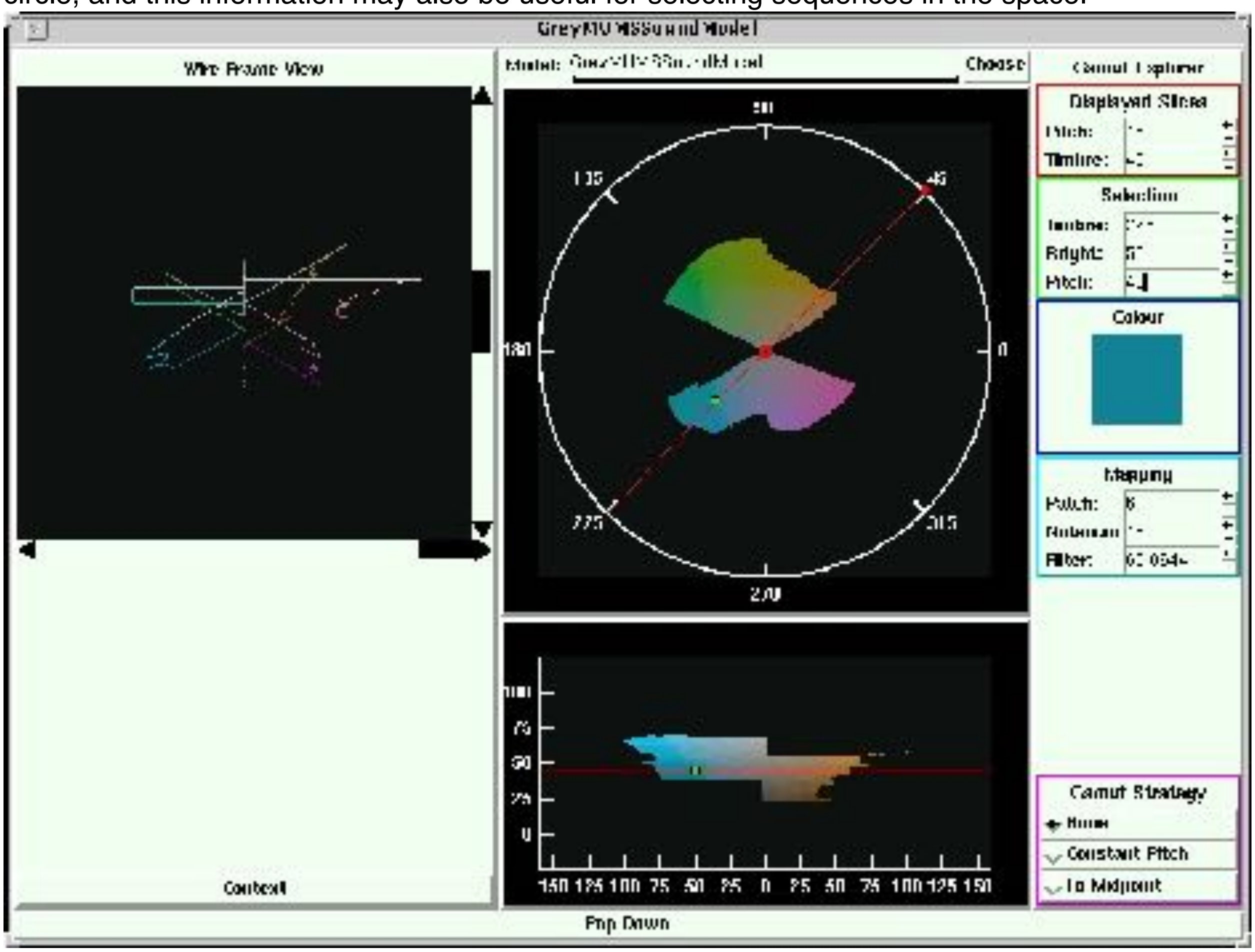
The abstract TBP Sound Space has some interesting properties which can be used to verify the internal behaviour of the concrete realisation. A vertical line should be a pitch scale with constant brightness and timbre. The pitch scale and timbre constancy were checked both numerically and by listening. A radial line should be a brightness scale of constant pitch and timbre. This was also checked numerically and by listening. A horizontal slice through the space should have constant pitch but will vary in timbre and brightness. A representative subset of points was checked numerically and by listening. A circle of constant radius should vary in timbre but remain constant in brightness and pitch. The change in timbre and constancy in pitch were checked numerically and by listening. The testing of brightness constancy by listening proved to be a more difficult problem. The method used was to repeat the sound sequence at a rate of 1 sound per second for 50 cycles. After a few repetitions a "tune" becomes quite distinct (musically known as a Klangfarbenmelodie), even though all pitches are equal. When the brightnesses of the timbres are similar the tune becomes more difficult to perceive. It may be that the tune effect is related to the variation in the brightnesses of the sequence of timbres - a flatter more constant profile results in less variation making the tune more difficult to discern. This effect was used to check the brightness constancy for regularly spaced brightness contours, by listening for a "constant profile" Klangfarbenmelodie in each of them.

4.5. GAMUTEXPLORER

The GamutExplorer, shown in Figure 12, provides an interactive visualisation of the Concrete Sound Space. This tool was conceived to support interaction with perceptual colour spaces

(Robertson P.K., Hutchins M., Stevenson D., Barrass S., Gunn C., Smith D. 1994). It has been extended by using colours to visually represent sounds, and by adding an auditory interface. Different timbre categories are represented by different hue categories. Changes in timbre brightness are represented by changes in hue saturation. Changes in the pitch continuum are represented by changes in the lightness continuum. The top left diagram shows a 3D wireframe view of the skin which defines the boundary of the gamut of sounds in the space. The top right panel shows a horizontal slice through the space, and the lower right is a vertical slice. The user may explore the space by selecting different viewpoints in the wireframe view using scrollbars, and by selecting different slices in the slice views. Tapping with the left mouse button on a point in either of the slice views allows the sound to be heard.

The visualisation allows an understanding of how the different pitch and brightness ranges of each timbre interact with each other. In the horizontal slice, shown for pitch 46, it is immediately obvious that the flute and soprano sax are not available. Also note the changes in brightness around the timbre circle, for example the bassoon has a lower dynamic range in brightness at this pitch than its neighbours. The vertical slice shown in the figure is for timbre angle 45, which cuts through the cello and the English horn. The difference in pitch ranges of these timbres is immediately obvious, making the selection of a sequence which passes smoothly through the pitch intersection very easy. This view also allows a comparison of the brightness profiles. These diametrically opposite timbres are most dissimilar in terms of the structure of the timbre circle, and this information may also be useful for selecting sequences in the space.



[Figure 12. The GamutExplorer]

5. SUMMARY

This is an account of the design and construction of a perceptually organised sound space to support information representations. Four principles of information perception were proposed, based on an abstraction of the use of colour spaces to represent data in scientific visualisation. These principles - completeness, comprehensibility, consistency, and cohesiveness, match data structure with perceptual structure to improve the natural comprehension of data displays. The construction of a concrete instance of this organised sound space was carried out in four stages - the pedestal, the skin, the skeleton and the flesh. The raw material for the construction was the MUMS palette of musical instrument samples, which provides a reference for other researchers in the area of auditory perception. The pedestal consists of 8 equally spaced timbre steps organised in a circle by similarity. The skin is an arrangement of brightness profiles which defines the limits of dynamic range for variation in pitch and brightness at each timbre. The skeleton consists of a grid of equal differences in brightness and pitch for each timbre. These grids are welded together down a central axis and radiate outward like segments of a mandarin. The flesh is a continuous medium moulded to the skeleton and skin to model the behaviour of the overall space. The resulting sculpture has the property that there is a relationship between distance and the likelihood of perceptual grouping between points in the space. A vertical line is a pitch scale, and may be used to naturally represent continuous data types. A radial line is a scale of equal brightness increments for a timbre, and may also be used for continuous data types. A circle of constant radius is a contour of constant brightness across the range of timbres which can be used to represent categorical data types. These properties are a rich area for further experiment with data mappings and simultaneous and sequential presentations of the selections.

The auditory display methodology proposed here focuses on the perceptual relationships between organised groups of sounds, and the consideration of representations for a general range of data types. This approach is significantly different from other auditory display methods which do not consider data types or perceptual structure at all.

6. ACKNOWLEDGEMENT

This research was supported by a scholarship from the CSIRO Division of Information Technology. I would like to thank my supervisors, Dr. Phil Robertson of the CSIRO Division of Information Technology and Mr. David Worrall of the Australian Centre for Arts and Technology, for their support. I would also like to thank Dr. Don Bone, Dr. Ken Tsui and Ms. Danielle Landy of the CSIRO Division of Information Technology for their reviews of this paper.

7. BIBLIOGRAPHY

[1] Barnard P. 1991. Bridging Between Basic Theories and the Artifacts of Human-Computer Interaction, in Carroll J.M. (ed), Designing Interaction : Psychology at the Human-Computer Interface, Cambridge University Press.

[2] Barrass S. 1994a. A naturally ordered geometric model of sound inspired by colour theory. Proceedings of Synaesthetica 94, Australian Centre for Arts and Technology, Canberra.

[3] Barrass S. 1994b. A perceptual framework for the auditory display of scientific data. Proceedings of ICAD `94, Santa Fe, New Mexico.

[4] Bone, D. 1993. Adaptive Colour-printer Modelling Using Regularized Linear Splines. Proceedings of Device Independent Color Imaging, IS&T/SPIE Symposium on Electronic Imaging Science and Technology 93.
[5] Bregman A.S. 1990. Auditory Scene Analysis : The Perceptual Organisation of Sound, Cambridge: The MIT Press.
[6] Ehresman D. & Wessel D. 1978. Perception of Timbral Analogies, IRCAM, Paris : Centre Georges Pompidou.
[7] Grey J.M. 1975. Exploration of Musical Timbre, Phd. Thesis, CCRMA Dept. of Music, Stanford University, Report No. STAN-M-2.
[8] Kramer G. 1994. Auditory Display: Sonification, Audification and Auditory Interfaces, SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII, Reading : Addison-Wesley Publishing Company.
[9] Norman D.A. 1991. Cognitive Artifacts, in Carroll J.M. (ed), Designing Interaction : Psychology at the Human-Computer Interface, Cambridge University Press.
[10] Opolko F. and Wapnick J. 1995. McGill University Master Samples, http://lecaine.music.mcgill.ca/MUMS/Html/Mums.html.
[11] Pachella R.G. Somers P. and Hardzinski M. 1981. A Psychophysical Approach to Dimensional Integrity. In Gett D.J and Howard Jr .J.H. (eds.). Auditory and Visual Pattern Recognition. Hillsdale : Lawrence Erlbaum Assocs.
[12] Pollack I. and Ficks L. 1954. Information of Elementary Multidimensioal Auditory Displays. Journal of the Acoustical Society of America, Vol 26, 1550-158.
[13] Plomp, R. 1976. Aspects of Tone Sensation. London : Academic Press Inc.
[14] Robertson P.K. 1988. Visualizing Color Gamuts: A User Interface for the Effective Use of Perceptual Color Spaces in Data Displays, IEEE Computer Graphics and Applications Vol. 8 Num. 5, 1988, 50-64.
[15] Robertson P.K., Hutchins M., Stevenson D., Barrass S., Gunn C., Smith D. 1994. Mapping data into colour gamuts : using interaction to increase usability and reduce complexity, Computers & Graphics, Vol 18, No 5, 653-665.
[16] Rogowitz B.E. and Treinish L.A. 1993. An Architecture for Rule Based Visualization, in Proceedings of IEEE Visualization '93, San Jose, California..
[17] Sandell G.J. 1995. SHARC Timbre Database, http://www.parmly.luc.edu/sandell/sharc/SHARC.homepage.html.
[18] Van Noorden. 1977. Minimum differences of level and frequency for perceptual fission of tone sequences ABAB. Journal of the Acoustical Society of America, 61, 1041-1045.
[19] Ware C. 1993. The Foundations of Experimental Semiotics: a Theory of Sensory and Conventional Representation, Journal of Visual Languages and Computing, Vol. 4, Academic Press.
[20] Wiersma W. 1991. Research Methods in Education, Sydney: Allyn and Bacon.
[21] Von Bismarck G. 1974. Sharpness as an Attribute of the Timbre of Steady Sounds, Acustica, Vol. 30, pp 159.
[22] Zwicker E. and Fastl H. 1990. Psychoacoustics Facts and Models, Berlin : Springer-Verlag.