Proceedings of the Institute of Acoustics

SYNTHESIS OF AUDIO WAVEFORMS FOR THE MEASUREMENT OF PERCEIVED ORGAN ENSEMBLE IN VARIED ACOUSTICS

L.K. Comerford (1), P.J. Comerford (1) and A.I. Katrami (2)

(1) Microcomputer Music Research Unit (Department of Computing), University of Bradford, UK
(2) Electronic Imaging and Media Communications Unit, University of Bradford, UK

1. INTRODUCTION

This paper discusses the use of digital synthesis to produce precisely controlled variations in audio waveforms for use in the study of the perception of ensemble in organ tone.

2. THE ORGAN AND ENSEMBLE

An organ is an instrument in which a number of sounds can be controlled simultaneously by a single player. Thus, as with an orchestral chorus, sounds of different pitches with a wide range of tonal characteristics can all be active together; this is known as "ensemble". Characteristics which are individual to separate sound sources in the ensemble can be used by the auditory system as cues to differentiate one sound from another. This promotes the perception of the presence of the separate sounds even when these are heard concurrently. The perceived degree of ensemble increases as the sound complex is perceived to be comprised of more musical resources [Comerford and Comerford 1995].

3. THE AUDITORY BACKGROUND TO ENSEMBLE PERCEPTION

In order to examine the perceived degree of ensemble in organ tone, it is first necessary to establish how the auditory system processes simultaneous sounds. When sources sound concurrently, two opposing tendencies are at work : the tendency for sounds to be perceived to mix together into one (or for some to mask the presence of others), and the opposite tendency, for the components of the ensemble to be perceived as separate entities. These tendencies are conveniently termed by McAdams [1982] as *fusion* (that is, of multiple sources into one) and *parsing* (that is, separation of sounds in an ensemble). When a number of complex sounds are active simultaneously, both fusion and parsing will be taking place to some extent.

3.1 Fusion

There are two different types of fusion:

3.1.1 note fusion

The human auditory system generally perceives the components of sound which emanated from the same source or from the same initiation event as being linked into a single entity. For example, when the key for middle C is struck on a piano, we hear a piano note of middle C and not a series of individual partials. In other words,

"we hear a unitary experience of a complex sound, to which we ascribe a pitch." [Moore 1977]

A number of factors give rise to this type of fusion of a group of partials into a single note, and as Bregman [1993] points out, the judgement of perceptual fusion depends upon a combination of these factors. These include:

3.1.1.1 ***harmonicity of partials:*** When a complex tone reaches the ear, a number of different location on the basilar membrane are excited concurrently. A "central pitch processor" in the brain [Roederer 1975] is able to calculate that, if the locations excited are of invariant frequency-distance one from another, then they are harmonics of one fundamental and therefore could all belong to a single entity. Thus the components of a harmonic sound will tend to fuse perceptually into a single pitched entity by virtue of being harmonically related, through the operation of the central pitch processor.

3.1.1.2 ***envelope consistency:*** If a group of partials are all initiated by a single activation event - such as a hammer strike on a tubular bell - there may be features in their amplitude and frequency envelopes which promote the perception of their status as components of a single note. These are:
synchronous initiation: The initiation event is the temporal marker for the start of the attack period for all partials in a single note. This will be of most significance in sounds like percussion instruments where the initiation event tends to be very sharp and the differentials between attack rates of partials tend to be small. Handel [1989] records that if partials' onset asynchronies exceed 30-40ms, the partials will not be heard as a single tone.
related envelope tendency or "common fate": In some circumstances, the envelopes of all partials of a single note exhibit related envelope changes all displaying a certain tendency. For example, if a player increases the volume of a sustained blown instrument note, all partials of that note will increase in amplitude, even though the relationship between them may alter and more partials may be introduced. Again, in struck or plucked instruments, all partials in a single note will share the envelope tendency of decay, even though the decay rates and duration's of different partials may be widely varied.
Where a note is subject to small degrees of envelope modulation - from random fluctuations in amplitude and frequency caused by involuntary inconsistency in note generation activity, or from the deliberate application of vibrato or tremolo - the degree of coherence with which the component partials are affected by these modulations is reflected in the degree of note fusion ([McAdams 1982], [Roederer 1975], [Handel 1989]).

3.1.2 fusion of multiple sound sources (blend)
When multiple sources sound concurrently, some of their individuality is lost; this process is known as "blend". Contributory factors to blend are:

3.1.2.1 ***coincidence of partials:*** When the frequencies of partials of different simultaneous sounds coincide, these partials will be less distinguishable as arising from separate sources. With regard to organ tone, Mercer [1951] points out that
"(an organ chorus) includes, besides unison stops, those at sub-octave, octave and super-octave pitches, with certain "mutation" stops in tune with some of the natural harmonics of the unison note"
and adds that, for two simultaneously sounding organ pipes an octave apart to blend,
"adequate harmonic development of the lower-pitched pipe is the principal requisite; thus the harmonics of the higher-pitched pipe reinforce tones already existing." [Mercer 1953:]

3.1.2.2 ***amplitude masking:*** The spectra, synchrony and relative volumes of the sounds determine the degree of masking. If two tones start synchronously, the presence of the weaker tone cannot be detected by spectra alone if the volume levels of the two tones differ by more than about 20db, because the presence of the quieter tone will be completely masked by the louder tone [Rasch 1978].

3.2 Parsing

"The psychological study of complex tone superposition is still very incomplete. This applies particularly to the understanding of how the ear and brain are able to disentangle the "mess" of frequencies that belong to different simultaneous complex tones, so as to keep the sensations of those tones apart." [Roederer 1975]

Obviously, this "decomposition of the auditory mixture" [Bregman 1993] will depend partly on the strength of individual note fusions since if the components of a complex tone do not fuse there is no way to distinguish them from other concurrent sounds.

The parsing decisions made by the auditory system are informed by a combination of different factors, including familiarity with the sound sources and the degree of variance between the simultaneous sounds.

"If different factors promote contradictory groupings of the sounds, the winner will be the groupings with the most factors favouring it or the grouping that is favoured by the factors that the auditory system prefers to use." [Bregman 1993]

Factors contributing to parsing include:

3.2.1 onset asynchronies

"Acoustic components derived from independent environmental events tend not to start and stop at the same time." [Bregman 1993]

For this reason, onset asynchrony is a parsing cue because of the perceptual expectation that the asynchronous elements derive from different sound sources. (This is something like the opposite of the note fusion cue synchronous initiation, discussed above.)

The moment when a musical sound is first detected by a listener is known as its perceptual note onset (PNO). PNO is not necessarily coincident with the player's initiating action (for example, the time at which a player depresses a key), or with the start of the attack period (for example, the time at which a piano hammer strikes the strings of a piano note). This is because of both instrument action time, and detection time - since the presence of a sound can be obscured during the early (quiet) stages of its attack either by noise from other musical sources or by background noise.

There is a difference between actual asynchrony and perceived asynchrony. Even though asynchrony cannot be identified as such below 60-80ms, asynchrony above 30-40ms will prevent note fusion occurring between partials (see above); the same sensitivity applies to parsing between different notes. This is an important factor in counter-acting the blending effects of amplitude masking in an ensemble: Rasch [1979] found that, when two-sound stimuli are presented, even a very soft sound (-50db) can be detected in the presence of a loud sound, provided that the soft sound has temporal precedence of between 10 and 30ms. If the notes are actually synchronous, however, the amplitude difference cannot exceed 20db if the presence of both sounds is to be perceived.

Asynchronism in perceptual note onset has two main contributory causes:

3.2.1.1 ***varied player response:*** Complete and precise synchronism between players of different orchestral instruments is perhaps not a physical possibility in a musical situation, nor is it a musical necessity. It might be expected that notes of an organ chord, initiated as they are by a single player and operating through a coordinated organ action, would tend to be more synchronised than the notes of an orchestral

harmony involving multiple players. However, non-synchronous key depressions occur in playing organ chords, both voluntarily (for artistic and interpretative reasons) and involuntarily (because of mistakes, because of uneven key resistance and adjustment and because of unequal finger strength). Note-sequence speed can also influence the player response, since:
"faster tempo goes with more synchronisation, and slower tempo goes with less synchronisation" [Rasch 1988]

3.2.1.2 ***varied note attack characteristics*****:** The organ comprises stops with different note initiation mechanisms (for instance, edge tone for flues and reed vibration for reed stops). These differences give rise to varied attack *characteristics*, with varying patterns of differential partial growth, overshoot, and so on. There is also wide variety of attack *times* in organ tone, caused by using pipes of different sizes, varying between about 32' and 1/3" in speaking length, and in width between approximately 2' and 1/8", producing notes with fundamental frequencies in the range 16Hz to over 16KHz. Even if keys are depressed simultaneously, the time taken by the organ action to deliver air to pipes of different sizes and locations will vary. Attacks can vary in duration from around 25ms to more than half a second.

3.2.2 fundamental frequency differences

"Differences in fundamental frequency of two simultaneously presented sounds ... enhance segregation and identification." [McAdams 1989]
When complex tones are heard simultaneously, as when an organ ensemble is played, some points on the basilar membrane will be responding to elements in more than one of the simultaneous tones. Where the frequencies sounding fall into the same critical bands, the frequency resolution derived from the spatially-analytical capacity of the basilar membrane is reduced. Therefore, complex sounds which are more widely separated in fundamental frequency will parse better than simultaneous sounds which are close in fundamental frequency, which will overlap more.

3.2.3 dynamics

Rasch [1978] has shown that if frequency modulation is applied to one of two simultaneously sounding notes, then this note can be detected at reduced amplitude levels, levels at which its presence could not be detected if it were not modulated. Similarly, McAdams reports [1989] that, by imposing both periodic (vibrato-like) frequency change and smaller random frequency fluctuations, upon a group of partials within a wide-pitched tone mass, the perception of those modulated partials can be fused into a single entity which is distinct from the background tone mass. Using different modulations upon further sub-sets of partials in the remaining tone mass causes them to segregate into distinct entities too. This effect applies even when the partials in each subset are inharmonic.

3.2.4 binaural information

Sound reaches the brain via the listener's two (differently located) ears. This binaural information can give clues of direction and distance of each sound source, since - provided that the sound source is not equidistant from both ears - the sound will arrive at each ear a little separated in both time and loudness, and the sound spectra will be attenuated in the ear which is furthest from the source (as that ear will be in the "sound shadow" of the intervening head). Even if the sound source is theoretically equidistant from each ear, in practice the listener moves his or her head in order to help locate the source of the sound. Thus using this binaural information, varied spatial locations can be used to distinguish the presence of different instruments and different organ pipes sounding simultaneously.

As with all fusion and parsing cues, spatial location is used in combination with other cues and its effectiveness by be modified by context - for example, Bregman [1993] suggests that
"the [parsing strategy] that tries to group sounds by their spatial origins may not be effective in reverberant environments."

4. SOUND CHARACTERISTICS AND PERCEIVED DEGREE OF ENSEMBLE

Thus there are counter-balancing tendencies which prompt the auditory system to classify simultaneous sounds either as melded together or as separate entities. When the characteristics of organ tone are considered in the light of these tendencies, it becomes apparent that the perception of ensemble would be affected by a number of different organ sound characteristics, individually and in combination. These are:

4.1 Transients

The term "transient" may be summarised as describing the portion of the sound which occurs:
- at the start of a sound : before full volume and/or sustained state is reached - known as the "attack" transient;
- at the end of a sound : when the sustained state ends, or when a gradual decay is interrupted by a damping or other ending mechanism - known as the "decay" transient.

As the word implies, a "transient" is therefore a temporary passing state. It is often a period of extreme spectral disorder since partials can attack (and decay) at different and altering rates. During attack partials can overshoot their sustain state values, and there can also be prominent unpitched wind noise overshoot. Even when the *order* of partials' attack cannot be perceived, the disorder and unpredictability of the attack transient is significant because
"the auditory system will treat a sudden change of properties as the onset of a new event." [Bregman 1993]
For this reason the attack transient draws attention to the start of individual notes and is important in establishing note identity. Provided that it is not obscured by background or concurrent sounds [Handel 1989], an attack transient will emphasise the presence of a particular articulating sound source within a complex whole.
There will be a wide variety of transient characteristics within an organ ensemble, with attack and decay envelopes and duration's differing both between different stops and between different notes on the *same* stop.

4.2 Frequency Imprecision

Frequency imprecision arises in the organ from two main causes : imprecise tuning and wind supply variations.
Just as between different instruments in an orchestra, deviations from precise tuning occur between notes within an organ stop and between different organ stops. Tuning imprecision can be caused by imprecise tuning by the organ tuner, or by the pipe frequency altering after tuning because of physical movement of the tuning device or because of temperature change. Furthermore, the use of equal temperament inevitably implies the presence of some frequency dissonance, even if all note of all stops were tuned with precision to the equal tempered scale, because the note intervals in this temperament do not precisely coincide with harmonic intervals. When the partials produced by de-tuned notes or stops are close but not identical in frequency, this causes a "beating" effect of amplitude variations as the waveforms cancel and reinforce

one another as they move in and out of phase. If the tuning imprecision is very great, the organ is perceived as out of tune.

Frequency imprecision can also be caused by wind supply variations. In the sustained tone of an organ note, there are dynamic variations in both amplitude and frequency caused by small random fluctuations in the supply of wind to the pipes. (These are similar to the random fluctuations occurring in the tone of other instruments which produce sustained sound, which are caused by variations in the strength of blowing or of bowing). The effects of these fluctuations in the air supply will vary in different pipe sizes and types, causing instabilities of varied speed and depth, which alter during the life of the note.

As noted above, such envelope dynamics will accentuate the fusion of individual notes, which will in its turn promote parsing within the tone mass.

4.3 Frequency Spread

"The concept of a chorus is fundamental to an organ of any size; it represents the extension of the unison tone upwards and downwards by several octaves, including stops representing off-unison harmonics of the fundamental." [Mercer 1953]

The ear can more readily detect sounds as separate entities if they are widely differentiated in frequency, because this activates more locations on the basilar membrane [Roederer 1975], whereas if multiple unisons are played, a limited set of locations on the basilar membrane is more heavily stressed.

As with most characteristics of musical sound, changing one parameter in an acoustic instrument sound will change others. In this instance, increasing the frequency spread of a pipe organ ensemble will also introduce a greater range of attack characteristics, since in general higher notes and higher pitched stops will attack faster than lower pitched sounds in the same family. A wider variety of timbre will also be introduced, since even within the same family there will tend to be sounds with a greater number of partials on lower pitched notes.

The frequency spread of an ensemble is also closely connected to the partials spectra of the contributory sounds, and to the perceived loudness of the ensemble. These characteristics, which are changed as a concomitant effect of altering the frequency spread of an ensemble, are examined below.

4.4 Partials' Spectra

The term "spectral centroid" is used to indicate the frequency range in which the acoustic energy of a sound is concentrated.

The stops on an organ vary widely in the frequency, number and amplitude of partials of which they are composed. The spectral centroid, and hence the tone, produced from each pipe will depend upon the pipe shape (e.g. cylindrical, conical); whether it is open or closed at one end; its length in relation to its width, which affects harmonic development and volume [Roederer 1975] - wide pipes (described as "large scaled") suppress upper harmonics and have a "flutey" tone, whereas narrow pipes (termed "small scaled") have stronger higher harmonics, and in consequence a "stringy" or "bright" tone. Tone will also be influenced by the shape of the pipe mouth, lip, and ears, and by the treatment of the pipe (for example, nicking and leathering), as well as by the type and condition of wood and metal used in the pipe construction. In addition, the partials spectrum of a stop will alter across its compass, with lower notes generally having more partials than higher notes. The precise relation of one partial to another in *amplitude* - particularly low order partials - is at least as important in determining tonal quality as the *number* of partials in a tone. Higher harmonics may fall within the same critical band of the auditory system and may therefore not be perceived individually [Handel 1989]. The spectral centroid of an organ ensemble will itself be dependent upon the centroids of the stops which comprise it.

Hartmann [1982] comments that it is sounds with strong upper partials which are salient within a musical mass, citing the example of the oboe which is used for orchestral tuning, and from which a single note can be heard "cutting through" the mass of sound of the orchestra tuning up. On the other side of the same coin, Sandell [1989] says that sounds with *low* spectral centroid will blend more than those with stronger upper partials. (However, without a certain degree of harmonic development, blend will be inhibited.)
Where stops within an ensemble have a wide variety of partials' spectra, parsing is likely to be increased since an effect analogous with increased frequency spread will be created. Introducing varied spectra into an acoustic ensemble will also result in an increased variety of amplitude envelope characteristics, onset asynchronies, tuning deviations and so on.

4.5 Perceived Loudness
Adding stops of higher pitch to build up an organ chorus is
"the normal way (which is unique to the pipe organ) of increasing the apparent loudness of the sound by artificially reinforcing the upper harmonic structure." [Padgham 1986]
Thus, to re-voice a pipe organ for a larger building,
"the 8' and 4' ranks are increased in power only a limited amount ... the principal increase in loudness is obtained by making the mixture larger and more complex and by the use of more powerful reeds." [Rienstra 1957]
It is analogous with the relative increase in the strengths of upper partials which occurs on other musical instruments when notes are played louder - such as by blowing louder on the trumpet or by bowing more vigorously on the violin or by striking the keys faster on the piano or by hitting a percussion instrument harder. It is true also of speech and of the singing voice. In all these cases there will be not just an increase in the amplitudes of partials, but a relatively larger increase in the amplitudes of higher frequency partials, and also an addition of new higher frequency partials.
The learned response of the brain is to associate an increase in higher frequency components with an increase in loudness, since this occurs in so many instances. Therefore, by adding more higher frequency partials by drawing in other stops of higher pitch or with higher spectral balance, or by opening the expression shutters to minimise high-frequency attenuation of enclosed stops, the player can create the perceptual illusion that the high frequencies of the original sound have been strengthened, thereby "implying" that the sound is being played louder. (The links between perceived loudness and both frequency spread and partials' spectra are thus very close in the organ.)

A further important influence upon the perception of organ ensemble, not in itself a sound characteristic, is the acoustic environment in which the sound is heard. Because of the bulky nature of the instrument, the sound of each organ is closely wedded to its individual environmental acoustic, so this needs to be taken into account in the measurement of perceived organ ensemble.

5. EXPERIMENTAL APPROACHES

To examine the contribution and interaction of these characteristics in influencing perceived degree of ensemble, and to be able to link a specific change or series of changes in a sound's content with an observed perceptual result, it is helpful to be able experimentally to control and alter the characteristics individually, and to combine them in defined ways. Three different experimental approaches are considered for this purpose:

5.1 Physical Adjustments to Acoustic Instruments followed by Perceptual Assessments of Sounds (i.e. a systematised version of traditional organ tone adjustment)

The traditional method of adjusting organ tone is to make physical alterations to the dimensions or shape of each pipe or to the wind supply to the pipe. This process is known as regulation and voicing. Voicing is

"a highly developed art but entirely empirical." [Mercer 1951]

Some changes made to pipes are reversible (for example, bending a metal pipe lip), although adjustments can probably not be measured with complete precision. Other changes (such as shaving off wood and cutting holes) are not fully reversible. The range of all such changes to pipes is relatively limited - for example, reed pipes cannot be converted into diapason pipes, and although a stop can be revoiced to sound louder such changes take a considerable time and the physical size of the pipes cannot be increased. The pipe organ regulator must also learn how the effects of the changes he makes will vary from one acoustic to another. This relates not only to the building as a whole but also to the "local" environmental effect, produced by the size, shape and disposition of the pipe chamber, the construction of the sound boards, and the presence of other pipes.

Commenting on attempts to assess the effects of building acoustic upon organ tone, Churcher [1965] says that ideally comparisons would be made

"on the same organ with the same registration in selected churches ... nothing but the acoustical environment would have been varied and any tonal changes observed could be confidently ascribed to it."

Writing of his work on pipe organ tone, (and before the availability of sufficiently sophisticated synthesis apparatus), he describes this as "impracticable". This is because, except for small portative instruments, it is not often possible to move a pipe organ from one building to another with no alteration, and as no two pipes are identical this makes the precise comparison of pipes in different acoustics very difficult. One possibility would be to remove a pipe, or set of pipes, record them in an anechoic chamber both before and after each of a series of voicing adjustments, and re-play these recordings in varied acoustics.

Despite there being much specialist and empirical wisdom about ways to achieve voicing effects, the physical adjustment method does not lend itself to easy experiments with variant voicings, nor to reversibility, or precise control, or exactly repeatable conditions, nor to comparison of voicings in different acoustics. (On a practical level, the authors have access to pipe organs in many locations - but not the permission to experimentally alter their voicing.)

There are examples of more systematic physical changes to pipe dimensions. For example, Kuhn [1940] built an organ pipe in which the languid and the lip positions, and the depth of the boot, were adjustable. He experimented with the effects upon the partials spectrum (and upon eddy formation at the upper lip) of altering the relative positions of these parts. He judged that the relative positions of the upper lip and languid were particularly significant in affecting the partials spectrum and the eddy formation. This was a short but painstaking study, and illustrates the importance of slight dimension changes. However, Kuhn's experimental apparatus was necessarily limited to one particular pipe of specific tonal type and pitch, and to certain adjustable parts in it, movable only in a specific way, so only a limited though interesting and useful set of measurements could be taken. (Also, there is no way of knowing whether pipes without adjustable parts do in fact behave in exactly this way.) The *perceptual* effects caused by the changes to the sound were not assessed.

5.2 Perceptual Assessments of Acoustic Instrument Sounds followed by Analysis

One could listen to many acoustic variants (for example, to notes on many different examples of diapason pipe stops), and assess each variant as producing a specific perceptual result. Each variant could then be analysed to display the difference in its content.

This would be a possible though cumbersome method of identifying the causes of a particular perceptual effect. However, it would be difficult to determine in this way which difference in content accounted for a perceptual change, since the changes in content that differentiate the variant tones one from another are likely to be complicated, and to involve a number of different factors which cannot be isolated (including the influence of different building acoustics on examples gathered from different locations). Furthermore, this approach does not allow the same organ sounds to be assessed in different acoustics (except by making recordings in an anechoic chamber as suggested above).

5.3 Synthesis of Audio Waveforms Followed by Perceptual Assessment of Sounds

In this approach, synthesised sounds are used, which vary in a known way. (The variation may be in the sound waveform content or in the environment in which it is heard.) An assessment is made of the difference in perceptual result of the sounds. Using this approach, the effect of a particular variation can be identified and, conversely, the contributory causes of perceptual differences can more easily be traced (although it must be recognised that perceptual effects can arise from *combinations* of variants).

Assuming that there is available simulation equipment of sufficient sophistication to allow synthesis and manipulation of high-quality sounds, this approach has the advantages that:

- the adjustments are precisely measurable.
- the adjustments are precisely repeatable, so they can be tested in a range of different musical situations and in a variety of acoustics.
- adjustments can be made to different characteristics which contribute to a tone, without the need to construct special test apparatus for each characteristic; similarly, adjustments can be made to many different tone types, without need to obtain or construct special test apparatus for each type.
- in some cases, adjustments can be made to characteristics in isolation, without effecting other characteristics of the sound (which is not possible in an acoustic instrument); on the other hand, different adjustments can be combined for assessment.
- synthesis adjustments can be used to examine experimental features not available on pipes (for example, putting a reed transient onto a flue steady state).

The method fulfils the criterion for methodology in the psychology of music, defined by Sloboda [1986] and summarised by McAdams [1987] as:

"synthesis of sounds by computer in order to have precise control over the way a sound varies and then to be able to relate these changes to variations in reported musical experience."

The validity of this approach and the advantages of using synthesised variants in musical psychoacoustic testing may be summarised thus:

"Experiments with electronically synthesized music can help bridge the gap between psychoacoustics and music. On the one hand, electronic synthesis can be used to create real music, music which will stand up to repeated listenings by musically sophisticated listeners. On the other hand, electronic synthesis offers a degree of precision which characterises psychoacoustic experiments. Therefore, electronic synthesis enables one to perform well-controlled experiments within a musical context." [Hartmann 1982]

The feasibility of using synthesised variants in this study depends upon the availability of a synthesis system able to produce organ syntheses of sufficient quality to trigger real musical judgements, and controllable with a sufficient degree of precision to allow adequate experimentation. The authors have access to The Bradford Musical Instrument Simulator (BMIS), a general purpose microprocessor controlled musical sound synthesis system, capable of the producing high quality sound, and of very precise sound control [Comerford 1993]. BMIS has been developed principally for the production of organ tone, and is a highly sophisticated system within its specialist field.

In view of the advantages of this experimental approach and the availability of a suitable synthesis system, this is the approach chosen for the current study, making use of the simulation facilities afforded by BMIS.

6. SYNTHESISED WAVEFORM MANIPULATIONS

The syntheses were needed for experiments to examine how changes in sound characteristics altered perceived degree of ensemble in organ tone, and to discover the influence upon the perceived ensemble when identical organ sounds were heard in different acoustics.
Except in the partials' spectra and frequency spread experiments, a three-stop diapason ensemble was used throughout, with stops pitched at 8', 4' and 2'. An intermanual coupler to pedal was also selected to extend the note range of the musical examples. The synthesised waveform manipulations used may be summarised as follows:

6.1 Transient complexity

Two variants were created. The first represented the "typical" perceptual complexity of transients in acoustic organ ensemble. Amplitude attack envelopes comprised two differential partial attack rates per sounding note, appropriate to each stop and to note pitch, and included harmonic and inharmonic partial overshoot and a burst of simulated air sound. The frequency envelope during attack incorporated a frequency shift. There were two differential partial decay rates per sounding note, as appropriate for each stop and note, and a frequency shift during decay.
The second variant was of "reduced complexity", and had simply two differential partial attack rates per sounding note, appropriate to each stop and to note pitch.

6.2 Frequency imprecision

Frequency imprecision syntheses included tuning deviations from the equal tempered scale applied randomly to each sounding note of each stop, and dynamic frequency instability.
Alternative tuning ranges were, for high tuning deviation, +/- 0.364%, and for low tuning deviation, +/- 0.065%. Dynamic frequency movement (instability) varied between +/- 0.004% (most stable) and +/- 0.017% (least stable).

6.3 Frequency spread

Three diapason ensembles were created, representing varying degrees of frequency spread. These comprised "high spread" (using one unison, one octave and one superoctave); "medium spread" (using 2 unisons and 1 octave); and "low spread" (using 3 unison stops). The frequency imprecision was held constant for all spreads (+/- 0.171%), but the transient characteristics and stability were appropriate to the stops' pitches and voicings.

6.4 Overall amplitude

In order to be able to separate the effects of overall amplitude from the elements of perceived loudness in their influence upon the perceived degree of ensemble, syntheses were produced which varied only in overall amplitude without any change to the waveform content. The "typical" amplitude, at which all other syntheses were presented, was that judged to be appropriate in the given acoustic. Alternative syntheses were presented at 3.5db (average) higher than the "typical" amplitude.

Details of the partials' spectra syntheses are reported elsewhere [Comerford 1995].

7. ASSESSMENT OF SYNTHESIS APPROACH

The overall aim of the experimental work to which these synthesised waveforms contribute is to assess links between an organ sound's characteristics and its perceptual effects in terms of ensemble. An important secondary consideration is the need to be able to compare the effects of the same sound characteristics heard in different environmental acoustics. The synthesised waveforms enabled manipulation of different sound characteristics individually, in a way not possible in acoustic sound. This allowed, for example, for a variety of partials' spectra to be presented with identical amplitude envelopes, to exclude the influence of the changes in transient characteristic which usually accompany varied stop types. However, the flexibility available in the syntheses had to be tempered by the need to retain stimuli as like their acoustic counterparts as possible, in order to ensure their musical and perceptual validity. An example of this is given above in the description of the frequency spread syntheses.

Each synthesis was precise and repeatable. This meant that the different waveform syntheses could be presented identically for assessment in each acoustic, and also for assessment between the acoustics.

Given the importance of using valid musical stimuli [Comerford and Comerford 1995], an added advantage of using syntheses was the simplicity with which an event-sequencer could be used to record performance details. This meant that the note timings of real musical passages played by an organist could be repeated identically in successive experimental presentations. (It would not be impossible to do the same using an acoustic organ, but in practice the likelihood of a suitable sequencer being fitted to each pipe organ used for experiments would be remote.)

The results obtained from the experimental work are reported elsewhere ([Comerford and Comerford 1995], [Comerford 1995]); in essence, they show that changes in organ tonal characteristics produced similar perceptual results in both the environmental acoustics tested, but that when the sounds were compared directly between the acoustics, the more reverberant acoustic correlated with increased perceived ensemble in every case.

8. REFERENCES

BREGMAN, A.S. Auditory scene analysis : hearing in complex environments Chapter 2 in McADAMS, S. and BIGAND, E. ed *Thinking in sound : the cognitive psychology of human audition* 1993 Clarendon Oxford

COMERFORD, L.K. The perception of ensemble in organ sound. PhD Thesis Bradford 1995 (in preparation)

COMERFORD, L.K. and COMERFORD, P.J. "The best stop on the organ?" - environmental acoustic and the perception of organ ensemble. *Proceedings, CIARM95 2nd International Conference on Acoustics and Musical Research* 1995 Ferrara pp 375-380

COMERFORD, P.J. Simulating an organ with additive synthesis. *Computer Music Journal* 17:2 Summer 1993 pp55-65

CHURCHER, B.G. Organ tone and acoustical environment. *The Organ.* (1965) pp. 68-85

HANDEL, S. Listening: an introduction to the perception of auditory events. 1989

HARTMANN, W. Electronic music : a bridge between psychoacoustics and music. in CLYNES, M. ed *Music, mind and brain : the neuropsychology of music* 1982 Plenum New York pages 371-385

KUHN, K.T. Klangfarbe und wirbelform einer lippenpfeife in abhangigkeit von der bauweise. PhD thesis. Berlin 1940.

McADAMS, S. Spectral fusion and the creation of auditory images. in CLYNES, M. ed *Music, mind and brain : the neuropsychology of music* 1982 Plenum New York

McADAMS, S. Segregation of concurrent sounds. I: effects of frequency modulation coherence. *Journal of Acoustical Society of America* vol 86 no 6 December 1989 pp2148-2159

MERCER, D.M.A. The physics of the organ flue pipe. *American Journal of Physics* 1953 pp376-386

MERCER, D.M.A. The voicing of organ flue pipes. *Journal of Acoustical Society of America*. Vol 23, (1), January 1951, pp45-54

MOORE, B.C.J. *Introduction to the psychology of hearing*. Macmillan 1977

PADGHAM, C. The scaling of the timbre of the pipe organs. Acustica Vol 60 (3), 1986, pp189-204.

RASCH, R.A. The perception of simultaneous notes such as in polyphonic music. *Acustica* vol 40 1978 pp21-33

RASCH, R.A. Synchronisation in performed ensemble music. *Acustica* vol 13 1979 pp121-131

RASCH, R.A. Timing and synchronisation in ensemble performance. (Chapter 4) in SLOBODA, J. ed. *Generative processes in music* Clarendon Press, Oxford, 1988

RIENSTRA, A.R. Acoustical and organ design for church auditoriums. Journal of Acoustical Society of America Vol 29 (7) July 1957 pp783-788

ROEDERER, J.G. *Introduction to the physics and psychophysics of music*. Spring-Verlag 1975 2nd ed.

SLOBODA, J. Cognition and real music : the psychology of music comes of age. Psychol.Belg. 1986 XXVI-2, pp199-219