

画像から音風景を作成する試み*

1U-5

寺本 邦夫 大木 直人 阿部 圭一 岡田 謙一 松下 温†

慶應義塾大学‡

1 背景

最近、マルチメディアという言葉がいたるところで聞かれる。CD ROMの普及により、様々なマルチメディアソフトが作り出され、一般家庭でマルチメディアソフトを楽しむことができるようになった。現段階では、ユーザはソフトを視聴して楽しんでいるだけだが、今後、自分でマルチメディアを作成したいという要求が多く出てくると思われる。

しかし、マルチメディアの作成というものは非常に複雑で手間のかかるものである。確かに、その作業を支援するマルチメディアオーサリングツールが存在するが、それらを用いてもなお手間がかかり（ユーザはすべてのメディアを組み合わせなければならぬ）、またできることが限られている（音の利用は、BGMや、解説を読みあげる程度である）。今のままでは、本当のマルチメディアの普及はできないと考える。そこで、誰でも簡単に利用できるマルチメディア作成ツールが必要となってくる。マルチメディアを簡単に作成するには、ある一つのメディアを人間が提示すれば、それに合ったその他のメディアをコンピュータが自動的に作成してくれるのが望ましい。

そこで我々はそのような、画像、音、テキストの相互マルチメディア作成ツールを目指し、その一環として今回は画像からそれに合った音を作り出すシステムである Interactive Sound from Picture（仮称、以下ISP）を提案する。

2 提案

ISPは、写真や絵（風景）から、それに合った立体的な音風景を作り出し、音付きの画像を表示する

*The Idea of Creating Sound Scape from Picture

†K.Teramoto,N.Ohki,K.Abe,K.Okada,Y.Matsushita

‡Keio University

システムである。さらに、ユーザは作り出した音風景の中に入り込んで自由に動き回ることができる。

3 実装

ISPのシステムは、画像に音をつけていく立体音場作成部と、作成された音付きの画像を表示し、さらにユーザがその音空間内を自由に動き回れるシステム実行部からなる。

立体音場作成部

立体音場作成部では、あらかじめ用意された画像からそれらにみあった立体音場を作成する。しかし現時点では画像を解析し、音を自動的に生成するのは不可能であるので、ユーザは画像上で音源の場所とその音源の種類を決定しなければならない。まずユーザは画像上で音源の位置を決める、例えば、写真に滝があって、その滝の音を出したい時、滝の所でマウスをクリックする。そうするとメニューが現れるので（図1）、その中で音源の種類、音源の距離を決定する。そして、クリックしたマウスの位置と、入力した距離から二次元音空間上の音源の場所を決定し配置する。

また、音源には、波の音、川のせせらぎ、木のざわめき、犬の鳴き声、寺の鐘の音など様々なものが考えられる。これら多くの音源を扱うためにメニューの中で音源の種類は階層化した（図2）。

システム実行部

システム実行部では、ユーザは、作り出した音場の中を自由に動き回ることができる。ユーザの位置は、顔を後ろから見たアイコンで表される（図3）。ユーザは前後左右の4方向に移動することができ、左右方向はアイコンも左右に動き、前方向はアイコ

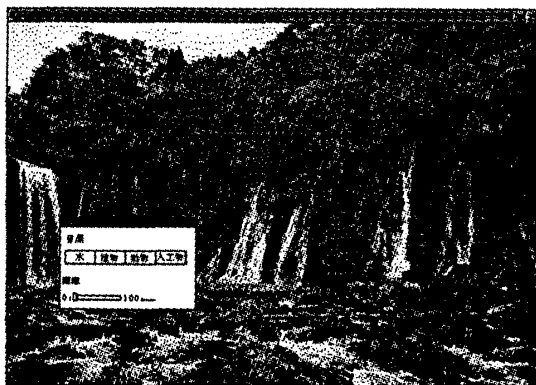


図 1: 立体音場作成部



図 3: システム実行部

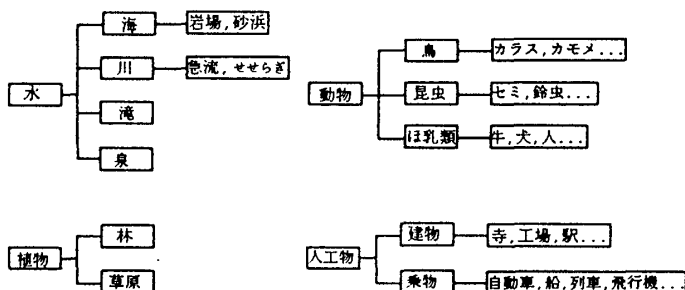


図 2: 音源の種類の階層化

ンが小さくなり、後ろ方向はアイコンが大きくなることにより表わされる。また音場の生成には、リアルタイムにユーザの動きに応じた音像定位を作り出すシステムである ISF (Interface with Interactive Sound Field) を用いた [1][2]。

4 実用例

現在のところ試作システムでは、マルチメディア旅行パンフレットを作った。これを用いると、写真に入り込んで仮想的に旅行の下見をしながら旅行のプランを立てることができる。また、写真は実際の場所に対応してリンクが張っており、ユーザは写真間を簡単に移動できる。例えば湖に近づけば、湖のアップの写真が現れるのである。

5 課題

システムは現段階では試作段階である。

音源はそれぞれ一種類につき一つの音しかなく、例えば、滝を選ぶといつも同じ滝の音が鳴っている。しかし、滝の音にしても様々な音があり、それらの場合によって使い分ける必要があると考える。例えばある写真一つとっても見る人によって感じ方は異なるだろう。ある人はそれを「もの悲しい」と思い、他の人は「いきいきとした」と思うかも知れない。ましてや、それが自分が旅行した時にとった写真なら、その景色を見た時の感情というものがあるだろう。今後は、心理学を用い、感性語も入力することによって、同じ写真でも使う人によって異なった音風景が作り出されるようにするつもりである。

また、システム実行部でのユーザインタフェースの改良も行なうつもりである。

さらに、背景でも述べたように、テキスト、音、画像の相互マルチメディア作成ツールに拡張していくつもりである。

参考文献

- [1] 大木, 亀倉, 阿部, 岡田, 松下, "人工現実感を用いた情報検索", 情報処理学会情報メディア研究会報告 IM11-1, 1993
- [2] 阿部, 大木, 亀倉, 岡田, 松下, "インタラクティブな臨場感を持った音場インタフェース", 情報処理学会全国大会, 1992