

人間の感性を考慮し、

5D-3

風景描写文から風景画像を作り出す試み*

大久保 達真 大木 直人 寺本 邦夫 岡田 謙一 松下 温†

慶応義塾大学‡

1 序論

近年の情報機器の高性能化や低価格化、WWW (World Wide Web) の普及により、一般ユーザでも簡単にマルチメディア情報を創作できる制作環境の実現が期待されている。

一方、現在のマルチメディア情報制作は、複雑で時間のかかる退屈な作業であると報告されている。この問題を解決し、一般ユーザでも、簡単にマルチメディア情報を制作できる環境として、1) 扱いやすい検索キー、2) ユーザの感性を考慮した検索、という特徴をもつ必要がある。

以上の要求を満たす制作環境を実現するため、画像や音声を部品ごとに管理し、ユーザのイメージを入力することにより、イメージにあった部品を検索し合成する技術を提案する。本論文では、とくに風景描写文から、風景画像と音風景を作り出すシステムについての実装と実行結果について述べる。これにより、一般ユーザでも、すぐに、簡単に、マルチメディア情報を創り出すことができる。

2 風景画像への感性の反映

システムの出力する風景画像と音風景は、感性を反映したものでなければならないが、音風景への感性の反映は、我々の以前の研究で行った [2]。風景画像に感性を反映させる要因としては、構図、色、物体のイメージが考えられるが、本論文においては、主に構図と感性の関係に注目した。

構図と感性に関する知識は、長年にわたり画家など美術家により培われ、現代では体系化されている。本研究では、線や地平線が人間の感性に与える印象に関する知識を利用した [1]。例えば、水平線（広々とした、穏やかな）、垂直線（緊

張感）、斜線（動的な）、直線（固い）、曲線（柔らかな）などがあり、太さや長さはその線の性質を強調する。また、直線の組合せで、三角形（正方向）となるものは、安定した印象を与える。この知識を利用するために、あらかじめ画像データベースを構築する際に、各画像パーツの形を解析し、それぞれの画像パーツの直線度、曲線度、水平線度、垂直線度、平均斜度、三角形成度を求め、それらの値をもとに、各画像パーツのメンバシップ値を決定しデータベースに記述しておく。このメンバシップ値は、画像パーツ検索時に参照される。

また、風景画像において、空が広いと「広々とした」、空が狭いと「圧迫感」といった印象を与える。これは、画像における地平線の位置の高さと関係している。よって、ユーザの入力した感性語によって、地平線の高さを調節することにする。

3 システムの構築

我々の構築したシステムの概要を、図1に示す。まず、システムは、風景描写文を理解する必要があるが、自然言語理解の分野は、まだ研究途上であるため、風景描写文から、風景の合成に関係する単語を抽出する方法を採用した。抽出された単語は、山や鳥など部品の名前に関するもの（オブジェクト）、季節、時間、場所などの条件（客観条件）、ユーザの出したいイメージや雰囲気に関するもの（感性語）に分類される。

次に、ユーザに入力されていないが、その風景にあってしかるべきオブジェクトを補完する。例えば、港の風景には、カモメの鳴き声や灯台、漁船などは、必ず存在する。このようなオブジェクトは、ユーザの入力がなくても、システムにより自動的に付加される。補完は、入力されたオブジェクト一つ一つについて行ない、閾値を決定してその閾値以上の補完度を持っているオブジェクトを補完する。閾値は、ユーザ側で変更する事ができ、

*Generation of Landscape Reflecting Human Feeling

†Tatsuma Ohkubo, Naoto Ohki, Kunio Teramoto, Kenichi Okada, Yutaka Matsushita

‡Keio University

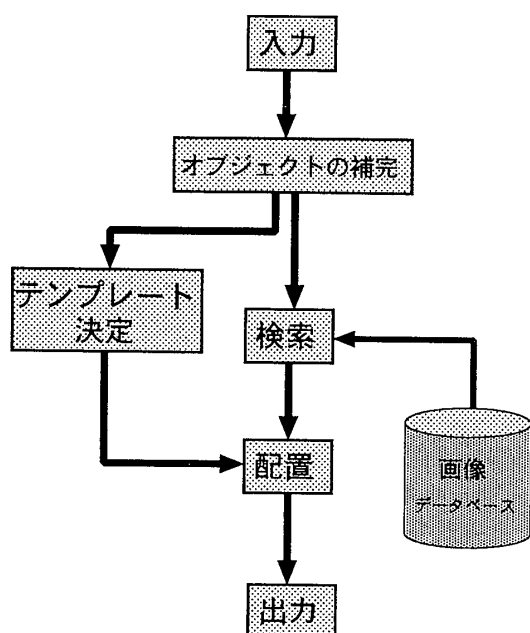


図 1: システムの概要

閾値を高くすれば、文中の語句に忠実な風景画像が作成され、閾値を低くすれば想像力豊かな風景画像が作り出されるのである。

抽出されたオブジェクト群にふさわしいテンプレートを選ぶ。テンプレートには、おもに、画像の背景とその背景の中で、画像パーツを配置するにふさわしい位置が、記述されている。テンプレートは検索キーワードを持っており、オブジェクト群とキーワードとのマッチングを行ない、適切だと思われるテンプレートを選択する。

一枚の画像と音風景に合成するにあたって実際に風景画像と音風景に出力される画像パーツと音パーツは、すべての感性語と客観条件に対して、先に述べた構図を反映させたメンバシップ値を持っている。オブジェクトに含まれる画像パーツのうち、もっともメンバシップ値の平均が高いパーツが、実際に合成されるパーツとして選択される。

選択された画像パーツを合成するため、まず、地平線の高さを決める。「広々とした」と似たような感性語が入力されると地平線を下げ、「圧迫感」に似た言葉が入力されると地平線を上げる。次に、それぞれのパーツをテンプレートの情報をもとに配置する。ただし、オブジェクトは、どのオブジェクトの上に置けるか、と言う情報を持っており、(例

えば牛なら地面の上、船なら海の上) もしその位置に置けないならその近辺を探す。

音パーツの情報は、画像パーツの配置された位置とともに、我々が以前開発したシステム OKeS [2] に送られ、配置され音風景が構築される。

4 システムの実装と実行結果

本システムは、SUN SPARC STATION の X ウィンドウシステム上で、Motif を用いて開発した。サウンドインタフェースは、OKeS を利用した。

実装したシステムの入力インタフェースを用いて、ユーザは、テキスト入力スペースに自由に風景描写文を入力できる。出力インタフェースの例として、「のどかな草原に牛がいます」という入力に対する実行結果を図 2 に示す。

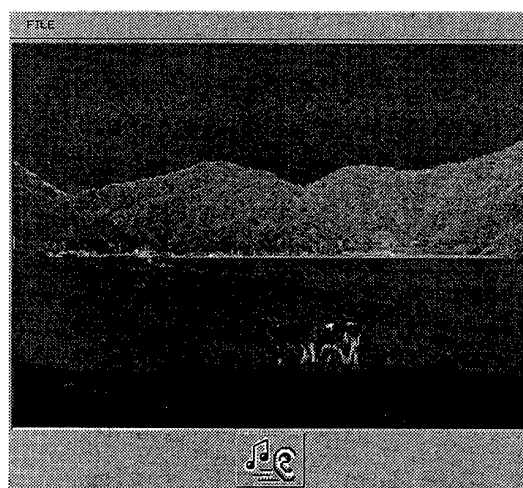


図 2: 入力: のどかな草原に牛がいます

文中に「空」「山」「地面」という単語は存在しないが、システムでそれらが必要であると判断され出力されている。また、地平線の高さは、空が広く現われていて「のどかな」感じを感じさせる。

参考文献

- [1] 視覚デザイン研究所: 構図エッセンス, 視覚デザイン研究所 (1983)
- [2] 寺本, 大木, 阿部, 内田, 岡田, 松下: 人間の感情を考慮した音風景の構築, 情報処理学会全国大会, 第 50 会, 4-327,(1995)