

平成15年度 情報工学コース卒業研究報告要旨

大西研究室	氏名	中村 知樹
卒業研究題目	画像・音声情報によるユーザの顔姿勢推定を利用したカメラの制御	

背景と目的

近年,端末間コミュニケーション手段としてカメラやマイクを設置して映像や音声を伝送することが行われているが,その際,奥行き情報としての姿勢の利用は多くはない.ユーザが端末に表示されている映像に対して,左右から覗き込むような姿勢奥行き情報変化から,対象の映像の側面を観察したいという意思を読み取り,側面を映し出すことができれば,端末間の通信により新しい表現を得ることができる.

本研究では,端末の前でユーザが姿勢を変化させた時に,ユーザの姿勢奥行き情報を推定するために顔画像と音声を利用したインタフェースを実装した.これを用いたカメラの制御の例を説明する.

システム概要

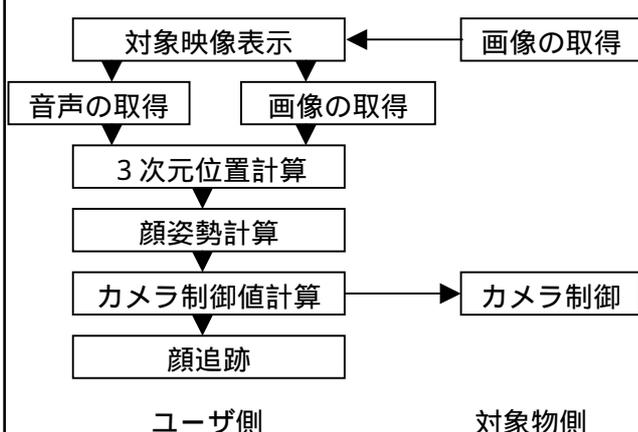


図1 システムの流れ

対象を撮影したカメラの画像を取得しユーザに提示する.そして,それを観察するユーザを2台のカメラで撮影し,さらに4台のマイクロフォンでユーザの音声を取得する.2台のカメラで撮影した画像は色情報(肌色)を利用して顔領域を抽出した後,得られた顔領域内を色情報(黒色)を用いて目領域を抽出した.目領域から三角測量法を用いて目の3次元位置を求める.

また,マイクで取得した音声を元に,各マイク間の音波到達時間差を相互相関法により求め,その情報を元に逐次近似法を用いて口の3次元位置情報を求める.

次に,画像による目の3次元位置と音声による口の3次元位置を利用して,ユーザの顔姿勢を各顔部位間のベクトルの外積から取得する.

顔姿勢変化からユーザの観察したい方向を推定し,対象を映すカメラを制御する.

実験

音声による口の3次元位置推定の精度の測定,顔姿勢の精度の測定,ユーザが対象物カメラを操作して目標物をディスプレイの中心に捕らえられるかどうかの実験を行った.

口の3次元位置推定の精度は,水平,垂直方向では誤差2cm程度,奥行き方向では,誤差5cm程度であった.

目標物を見つける課題を3回行い,画像の中央に目標物を捕らえるのに適切な時間で3回とも中心に捕らえることができ,システムのより自然なインタフェースの可能性を示すことができた.