

Roteiro

#####

Motivação

#####

- Computação Ubíqua

A Computação Ubíqua vem sendo tema de diversas pesquisas desde o início dos anos 90.

Mark Weiser propôs que a computação devia ser algo invisível, exigindo o mínimo de esforço possível permitindo ao usuário ter mais foco na tarefa do que na ferramenta.

Dentre todo esse cenário é fundamental os ambiente indigentes. Tais ambientes, também chamados de smart spaces, interagem com seus usuários auxiliando-os de modo proativo e transparente.

#####

- Ambientes Inteligentes

Ambientes são compostos por usuários mais dispositivos mais aplicações. Essas aplicações são softwares que provem serviços de modo a facilitar as tarefas dos usuários no ambiente.

Neste cenário, camadas de softwares denominadas de middlewares são responsáveis por orquestrar a interação e troca de informações entre aplicações, usuários e dispositivos.

Por exemplo: quais os dispositivos presentes no ambientes e os serviços disponibilizados, quais funcionalidade as aplicações provêm e de quais serviços necessita.

#####

- Informações de Contexto

A inteligência presente em um smart space é fruto de

aplicações que levam em consideração as informações de contexto do ambiente, isto é, informações sobre os usuários e os dispositivos.

São exemplos de informações de contexto:

- Temperatura do ambiente;
- Luminosidade;
- Informações dos usuários: número de usuários, frequência de entrada e saída, quem são eles, onde estão e o que estão fazendo;

Sem tais informações as aplicações não tem inteligência suficiente para interagir com os usuários de forma proativa. Essas informações são disponibilizadas pelo middleware às aplicações.

#####

- Localização e identificação (Onde são e Quem são)

Dentre os inúmeros dados de contexto que podem ser extraídos do smart space destacam-se a localização e identidade do usuário, que são fundamentais para a implementação de serviços proativos e personalizados.

A obtenção de informações acerca da localização e identificação dos usuários torna-se um desafio tendo em vista a dinamicidade do ambiente, onde usuários entram e saem a todo o momento, além de interagirem entre si e com diversos dispositivos.

Com a utilização de alguns sensores é possível obter dados sobre os usuários, que podem ser utilizados em sistemas de rastreamento, localização e identificação.

#####

Objetivo do trabalho

O objetivo desse trabalho foi o desenvolvimento de um sistema de rastreamento, localização, identificação de usuários em um ambiente inteligente que nos denominamos de TRUE.

TRUE é um acrônimo e vem de Tracking and Recognizing User

in the Environment.

O Sistema TRUE utiliza o sensor Kinect da Microsoft como dispositivo de entrada. Ele fornece imagens de cor e profundidade do ambiente.

No caso específico, o foco do nosso trabalho foi prover essas informações ao Middleware uOS.

O Middleware uOS é um projeto do grupo de pesquisa UnBiquitous da Universidade de Brasília que já existe desde 2007 coordenado pela Profª Carla e Prof Jacobi cujo foco está na adaptabilidade de serviços em um ambiente de computação ubíqua, de modo que os serviços dos dispositivos presentes no ambiente possam ser oferecidos e compartilhados.

#####

Identificação

Hoje em dia, várias técnicas de reconhecimento biométrico por meio da face, íris, voz, e outras, vêm sendo estudadas e utilizadas em sistemas de reconhecimento automático. Dentre estas, o reconhecimento por meio da face se destaca pois sua aquisição é realizada de maneira fácil e não intrusiva, tornando-a ideal para ser utilizada em um ambiente inteligente.

O reconhecimento facial tem suas dificuldades. Como:

- iluminação;
- ângulos;
- poses;
- expressões faciais;
- maquiagem e outros.

Basicamente, o processo de reconhecimento facial pode ser dividido em duas tarefas principais:

1. Detecção de faces em imagens;
2. Reconhecimento das faces encontradas.

#####

Identificação - detecção

A detecção de faces é definida como o processo que determina a existência ou não de faces em uma imagem e uma vez encontrada alguma face, sua localização deve ser apontada através de um enquadramento ou coordenada.

Atualmente, já existem diferentes métodos/técnicas de detecção de faces.

O método Viola-Jones é um dos métodos mais utilizados e possui uma alta acurácia permitindo a detecção de faces em tempo real. Este método pode ser utilizado para construir uma abordagem de detecção facial rápida e eficaz utilizando apenas imagens em tons de cinza. Apesar da simplicidade obtém altas taxas de detecção. O método é implementado pela biblioteca OpenCV.

#####

Identificação - reconhecimento

Na etapa de reconhecimento, as faces detectadas, serão comparadas com um banco de dados de faces conhecidas.

O Eigenfaces trata-se de uma técnica bastante satisfatória quando utilizada sobre uma base de dados relativamente grande. Ela usa o PCA(Principal Component Analysis) para extrair das imagens suas principais características e, partindo delas, realizar o reconhecimento das imagens utilizando um número bastante reduzido de cálculos.

#####

Localização

Para localizar um usuário nós baseamos em imagens de profundidade, onde cada pixel da imagem é atrelado a um valor de distância relativa ao sensor.

Para obter tais imagens tais imagens existem dois métodos o TOF - Time of Flight e a Luz Estruturada. Sendo a Luz Estruturada mais acessível que o TOF.

A Luz Estruturada consiste em adicionar uma textura artificial na cena e com uma câmera infravermelha realizar a análise.

#####

Rastreamento

O rastreamento pode ser definido como o problema de estimar a trajetória de uma entidade (objeto ou pessoa) em um plano de imagem à medida em que se move na cena. Em um ambiente inteligente, o rastreamento de pessoas é uma das principais ferramentas para detectar novos usuários e serve como base para que outras informações possam ser coletadas.

Basicamente, o processo de rastreamento pode ser dividido em duas etapas:

- Detecção do objeto;
- Rastreamento do objeto detectado.

#####

Rastreamento - detecção

Todo método de rastreamento requer um mecanismo de detecção de entidades que pode ser realizada a cada frame (quadro) obtido ou na primeira vez que a entidade a ser rastreada entra no campo de visão. Algumas dessas técnicas são:

- Detector de pontos;
- Subtração de fundo;
- Segmentação.

Subtração de fundo: é um método popular para segmentação de movimento, especialmente nas situações em que o plano de fundo é relativamente estático. Ele detecta as regiões de movimento na imagem obtendo a diferença pixel a pixel entre o quadro corrente e o quadro referente ao plano de fundo.

#####

Rastreamento

Os métodos de rastreamento de entidades mais utilizados atualmente são:

- Rastreamento de pontos;
- Rastreamento de núcleo;
- Rastreamento de silhuetas.

Rastreamento de silhuetas: o rastreamento é feito estimando a região da entidade a cada quadro a partir das silhuetas geradas nos quadros anteriores. Dado os modelos de entidades, silhuetas são rastreadas por qualquer forma de correspondência ou evolução de contorno.

#####

Trabalhos Correlatos

Agora serão mostrados alguns projetos que focam rastreamento, identificação e localização de pessoas em um ambiente inteligente.

Alguns dos projetos estudados utilizem abordagens multimodais, abordagens que usam mais de um tipo de dado de entrada como dados audiovisuais, este trabalho tem como foco identificação.

#####

Projeto CHIL

O Projeto CHIL (Computers in the Human Interaction Loop) envolve uma rede de quinze laboratórios internacionais de pesquisa acadêmica e industrial.

Este projeto utiliza informações obtidas de diferentes algoritmos monomodais para criar um sistema de rastreamento e identificação multimodal a partir de dados audiovisuais das pessoas no ambiente.

Este projeto não prove dados de localização. A identificação é

feita por meio de reconhecimento facial.

A foto contém um exemplo de uma face extraída do contexto.

#####

Smart Flow

Implementação de um sistema multimodal que realiza detecção de movimento, rastreamento de pessoas, reconhecimento facial, localização baseada em áudio no middleware SmartFlow.

Identificação e rastreamento por meio de imagens e localização por meio de áudio.

#####

AVIARY E MICASA

Nesse trabalho foi desenvolvido uma nova arquitetura para ambientes inteligentes chamado DIVA. Esta nova arquitetura pode ser vista como uma rede inteligente de câmeras, que são controladas para prover diversas funcionalidades, tais como:

- Identificação de pessoas;
- Detecção de intrusos;
- Rastreamento de múltiplas pessoas.

O rastreamento é feito utilizando subtração do fundo e o reconhecimento facial é feito utilizando o método Eigenfaces.

#####

Sistema TRUE

É um sistema aberto de rastreamento, localização e identificação de pessoas em ambientes inteligentes gerenciados pelo middleware uOS.

É um sistema monomodal implementado em C++ que utiliza somente dados visuais

Ele utiliza o sensor Kinect da Microsoft como dispositivo de entrada.

O Sistema TRUE é dividido em quatro módulos principais:

- Módulo de Rastreamento;
- Módulo de Reconhecimento;
- Módulo de Registro;
- Módulo de Integração;

#####

Middleware uOS

O Middleware uOS possui foco na adaptabilidade de serviços em um ambiente de computação ubíqua, de modo que os serviços dos dispositivos presentes no ambiente possam ser oferecidos e compartilhados.

É uma implementação de um middleware para auxílio de aplicações para ambiente de computação ubíqua, cuja proposta é compatível com os conceitos apresentados pela DSOA.

#####

Módulo de Registro

O Módulo de Registro é responsável por cadastrar novos usuários no sistema e treiná-lo para também reconhecer esse novo usuário.

O registro consiste das seguintes etapas:

1. Obtenção das imagens do novo usuário.
2. Processamento das imagens, isto é, as imagens são convertidas em escala de cinza, novas imagens são criadas recortando a região da face encontrada. Em seguida, as imagens são redimensionadas e equalizadas criando assim padrões de

tamanho, brilho e contraste.

3. Armazenamento das imagens.
4. Treinamento do sistema para reconhecer esse usuário.

#####

Módulo de Reconhecimento

O Módulo de Reconhecimento é responsável pela identificação dos usuários no ambiente utilizando a face como característica biométrica. A face foi escolhida por permitir um reconhecimento não intrusivo

O reconhecimento facial é feito a partir de imagens de usuários que são repassadas pelo Módulo de Rastreamento. Estas imagens são compostas somente pela região em que o usuário se encontra.

Basicamente o reconhecimento facial possui as seguintes etapas:

1. Obtenção da imagem de entrada.
2. Pré-processamento da imagem.
3. Detecção facial na imagem.
4. Processamento da imagem criando um padrão a fim de aumentar a acurácia do reconhecimento.
5. Reconhecimento facial com a técnica Eigenfaces.
6. Retorno do nome do usuário com a face “mais parecida” e a confiança do reconhecimento.

#####

Módulo de Reconhecimento - Detecção Facial

A detecção facial foi desenvolvida utilizando o método Viola-Jones que é implementado pela biblioteca OpenCV (Open Source Computer Vision).

Basicamente, o processo de detecção facial procura por uma face em uma imagem pré-processada.

Ele necessita de um classificador e o Sistema TRUE utiliza um classificador em cascata treinado para detectar faces frontais em imagens.

As etapas da detecção são:

1. Leitura do classificador.
2. Obtenção da imagem de entrada pré-processada.
3. Detecção da face na imagem utilizando o classificador.
4. Retorno da região da face detectada.

#####

Módulo de Reconhecimento - Reconhecimento Facial

O reconhecimento facial foi desenvolvido utilizando a técnica Eigenfaces.

Esta é uma técnica bastante satisfatória quando utilizada sobre uma base de faces relativamente grande, permitindo ao sistema inferir, das imagens das faces suas principais características e, partindo delas, realizar o reconhecimento facial utilizando um número reduzido de cálculos.

A técnica é implementado pela biblioteca OpenCV.

A base é composta por um banco de faces de alunos voluntários de Ciência da Computação da Universidade de Brasília

E um banco de imagens de faces da Universidade de Cambridge: formado por imagens de faces de 40 pessoas diferentes. Para cada pessoa, existem 10 diferentes imagens obtidas em diferentes momentos, com diferentes condições de iluminação, diferentes expressões faciais.

<passível de remoção>

O reconhecimento facial inicia-se projetando a imagem no subespaço através do método PCA que reduz sua dimensionalidade.

Então, calcula-se a distância da imagem projetada a cada um dos eigenfaces obtidos na etapa de treinamento obtendo uma lista de distâncias.

Esta lista de distâncias é comparada com as lista de distâncias

de cada usuário encontrando o usuário cuja lista de distâncias é a mais similar.

#####

Módulo de Rastreamento

O Módulo de Rastreamento é responsável por rastrear os usuários no ambiente, determinar a sua localização física em relação ao sensor Kinect e gerenciar suas identidades.

Para realizar o rastreamento e localização dos usuários é utilizada a biblioteca OpenNI (Open Natural Interaction).

Utilizando as imagens de profundidade, a detecção e o rastreamento dos usuários são feitos utilizando subtração de fundo.

A representação das pessoas detectadas e rastreadas são feitas utilizando suas silhuetas.

Utilizando imagens de profundidade é possível calcular as coordenadas (x,y,z) de um plano cartesiano de três dimensões em que o ponto (0,0,0) corresponde a posição do Kinect,

Dessa forma, a posição de um usuário em relação ao sensor é estimada utilizando as coordenadas referente ao pixel que representa seu centro de massa geométrico. Sendo assim, ao fixar a posição do Kinect no ambiente, é possível estimar a localização de qualquer usuário rastreado em tempo real.

#####

Módulo de Integração

Este módulo é responsável pela integração do Sistema TRUE com o Middleware uOS. Através desta integração, os dados providos pelo Sistema TRUE são disponibilizadas ao ambiente gerenciado pelo middleware uOS por meio de serviços.

Para integrá-los foi desenvolvido um driver para o Middleware se comunicar com o Sistema TRUE. Este driver foi nomeado de UserDriver.

Os métodos do UserDriver podem ser divididos basicamente em três grupos:

- Métodos de Inicialização: métodos responsáveis pela inicialização do Sistema TRUE e pelo registro dos serviços que o driver possui e os eventos que pode gerar.
- Métodos Nativos: métodos declarados no driver que permitem acessar alguns métodos implementados pelo Sistema TRUE. Tais métodos permitem que o UserDriver tenha controle sobre o sistema, podendo iniciá-lo, encerrá-lo, retreiná-lo e, também, cadastrar e remover novos usuários a qualquer momento.
- Serviços e Eventos: são os métodos que implementam os serviços que o UserDriver disponibiliza às aplicações presentes no ambiente e os eventos que podem ser gerados.

#####

Módulo de Integração 2

O UserDriver disponibiliza às aplicações registradas no middleware os seguintes serviços:

- Consultas: as informações dos usuários no ambiente: através dessas consultas, as aplicações tem acesso ao nomes, emails, posições correntes e confiança do reconhecimento de todos os usuários presentes no ambiente.
- Cadastro: as aplicações podem cadastrar novos usuários fornecendo ao UserDriver o nome, o email e as imagens do novo usuário.
- Treino do sistema: após cadastrar novos usuários as aplicações podem retreinar o sistema para poder reconhecer o novo usuário cadastrado.
- Remoção: as aplicações podem remover usuários cadastrados fornecendo o email do usuário.

Além dos serviços, o UserDriver gera os seguintes eventos:

- Novo Usuário: evento gerado assim que um novo usuário foi detectado pelo Sistema TRUE.
- Usuário Perdido: evento gerado assim que um usuário deixou de ser rastreado pelo Sistema TRUE.
- Atualização dos dados do usuário: evento gerado a cada 5

segundos atualizando os dados de todos os usuários rastreados.

#####

Testes

Com intuito de identificar a acurácia e as limitações do Sistema TRUE perante seus requisitos foram realizados uma série de testes funcionais.

Os testes foram conduzidos no LAICO (Laboratório de Sistemas Integrados e COncorrentes).

#####

Testes - Rastreamento - Detecção

Os primeiros testes foram realizados para testar a detecção.

Os testes foram feitos simulando a entrada de um usuário na cena e analisando o momento em que o mesmo era detectado.

A posição do sensor e a trajetória do usuário são apresentadas na Figura.

Em todos os testes o usuário era detectado antes mesmo de entrar na área de visão do sistema por completo, como mostrado nas Figuras abaixo.

#####

Testes - Rastreamento - Oclusão

Também foram realizados testes com o objetivo de verificar a oclusão no rastreamento.

Nos casos de oclusão parcial o sistema se mostrou robusto como pode ser observado na Figura da direita.

Nos testes foi simulada também a situação de oclusão momentânea, ou seja, um usuário em movimento oculta outro por um curto período de tempo em razão da sua movimentação.

Nestes casos, o sistema TRUE perde o usuário rastreado, mas rapidamente é capaz de detectá-lo novamente, como mostrado na

sequência de imagens.

A oclusão era um problema esperado uma vez que o Sistema TRUE utiliza somente um sensor Kinect como dispositivo de entrada.

#####

Testes - Rastreamento – Interferências

<passível de remoção>

Durante os testes realizados com rastreamento foram observado alguns problemas quando o usuário rastreado interagiu com objetos do ambiente ou com outros usuários.

Na grande maioria das vezes em que o usuário interagiu com objetos, o Sistema TRUE considerou o objeto como sendo parte do usuário. Entretanto, esta situação não prejudicou a eficiência do sistema.

Por outro lado, os problemas com interação entre usuários foram mais raros, porém tiveram impacto maior. Tais problemas consistem em “interferências” que ocorreram em algumas situações de contato entre dois ou mais usuários.

#####

Testes - Rastreamento

Apesar dos problemas relatados, o rastreamento conseguiu, na maioria dos testes, atender às necessidades rastreando os diversos usuários no ambiente em suas atividades diárias, como mostrado na Figura.

#####

Testes - Localização

O Sistema TRUE obtém a localização dos usuários no ambiente por meio de coordenadas dos mesmos em relação ao Kinect. Para verificar a acurácia dessas coordenadas foram realizados alguns testes.

As coordenadas (x, y, z) obtidas pelo sistema são coordenadas de um plano cartesiano de três dimensões em que o ponto (0, 0, 0) corresponde a posição do Kinect.

Os testes foram realizados para as coordenada do eixo x e z.

Para testar a precisão das coordenadas do eixo z foi realizado o seguinte teste: um objeto (uma caixa de papelão) foi colocada em frente ao sensor a diferentes distâncias do mesmo como mostrado na Figura.

Para testar a precisão das coordenadas foi realizado o seguinte teste: um objeto (uma caixa de papelão) foi colocado a uma distância fixa de 3 metros do sensor e colocada em diferentes posições ao longo do eixo x, como mostrado na Figura.

#####

Testes - Localização - Eixo Z

Para cada distância, foram coletados 6 amostras informadas pelo sistema. A Tabela apresenta o comparativo entre as distâncias reais medidas manualmente e a média das distâncias aferidas pelo sistema.

Neste teste o menor erro obtido foi de 3,21mm e o maior foi de 111,75mm.

Para distâncias acima de 4,057m não foram obtidos resultados pois essa é a distância máxima de funcionamento do sensor.

#####

Testes - Localização - Eixo X

Para cada posição foram coletadas 6 amostras informadas pelo sistema.

A Tabela apresenta o comparativo entre as distâncias reais medidas manualmente e a média das distâncias aferidas pelo

sistema.

Neste caso o menor erro obtido foi de 27,19mm e o maior foi de 79,29mm.

Acredita-se que os erros apresentados não são significativos a ponto de comprometer o uso dos dados fornecidos.

#####

Testes - Identificação

Para análise dos resultados foi utilizada como ferramenta uma matriz de confusão construída a partir dos dados coletados do sistema.

Tal matriz contém em suas colunas os valores possíveis que se podem obter para cada usuário e as linhas os usuários que foram observados.

Os valores presentes na matriz mostram os resultados, em porcentagens, dos testes de reconhecimento.

Cada célula apresenta o percentual de vezes que o resultado (coluna) ocorreu durante os testes com o indivíduo (linha) sendo reconhecido.

Desta forma, a matriz permite a fácil visualização da confiança obtida no reconhecimento observando-se os valores obtidos na diagonal principal.

Para sumarizar os resultados expressos na matriz apresentamos três taxas calculadas:

1. Verdadeiro Positivo: quando o sistema identifica o usuário de maneira correta.
2. Verdadeiro Negativo: quando o sistema identifica o usuário de maneira errada.
3. Falso Negativo: quando o sistema não identifica um usuário cadastrado.

Os cenários de teste consistiram de duas etapas:

1. Primeira etapa: cada usuário foi cadastrado na base de dados. Após o cadastro foram adicionados os dados das 40 pessoas presentes no banco de faces da Universidade de Cambridge.

2. Segunda etapa: cada usuário, individualmente, foi posicionado dentro do campo de visão do sensor onde foram coletados os resultados dos primeiros 20 valores retornados pelo sistema.

Primeiramente foi construído um cenário de testes que não apresentou resultados satisfatórios, com isto foi construído um segundo cenário de teste com uma nova abordagem para a coleta de faces.

#####

Testes - Identificação - 1º cenário

No primeiro cenário de teste, os cadastros dos usuários foram feitos utilizando 10 imagens de faces de cada usuário, sendo 6 imagens frontais da face, 2 imagens da face ligeiramente rotacionada para direita e 2 imagens da face ligeiramente rotacionada para esquerda.

O cadastro foi realizado com essa estratégia com o objetivo de diminuir o impacto das variações de poses e ângulo no reconhecimento facial.

Os resultados obtidos dos processos de reconhecimento facial realizados para cada usuário são mostrados na matriz de confusão.

Como previsto, pode-se observar na matriz que os valores mais significativos estão aglomerados na diagonal principal. Alguns resultados obtidos foram satisfatórios com porcentagens maiores que 90%.

Por outro lado, foram obtidos alguns resultados com baixas porcentagens como 45%.

Estes últimos foram principalmente causados por problemas como pose e expressões faciais.

A iluminação não foi um problema, pois no ambiente de teste não existe influência da iluminação externa e a interna é controlada.

A Tabela apresenta as taxas calculadas analisando os valores da matriz.

Observa-se que a taxa de verdadeiro positivo (73,63%) foi uma

taxa baixa para um sistema de reconhecimento automático.

Neste conjunto de testes, a taxa de verdadeiro negativo foi bem superior à taxa de falso negativo mostrando que a principal deficiência do reconhecimento facial consiste na confusão entre os usuários cadastrados no sistema.

#####

Testes - Identificação - 2º cenário

Sob a hipótese do resultado ter sido influenciado pela baixa quantidade de faces e poses, construiu-se um segundo cenário de testes.

Neste a estratégia utilizada nos cadastros dos usuários foi diferente.

Os cadastros foram feitos utilizando 100 imagens das faces de cada usuário em diferentes ângulos, posições e expressões faciais.

No momento do cadastro, foi solicitado que os usuários movimentassem o rosto levemente para cima, para baixo, para esquerda e para direita de maneira aleatória e com diferentes expressões faciais.

Os resultados obtidos neste conjunto de testes se encontram na segunda matriz de confusão.

Houve um aumento na taxa de Verdadeiro Positivo de 73,63% para 82,5% e redução significativa na taxa de Verdadeiro Negativo de 17,27% para 6,0%, mostrando que os casos de confusões entre os usuários foram mais raros.

A melhora obtida comprova a hipótese que um número maior de poses e faces aprimoraram a capacidade do sistema de reconhecer seus usuários.

#####

Testes - Integração

A Figura mostra as mensagens geradas por uma aplicação de exemplo que consome os dados providos pelo sistema TRUE.

Essa aplicação informa, via Twitter, quem chegou e quem saiu

do laboratório e também alerta sobre algum usuário que está sem se mexer a muito tempo.

#####

Conclusão

De acordo com os testes realizados, o Sistema TRUE apresentou uma boa acurácia na identificação de seus usuários, cerca de 80%.

Somado a isso obtivemos também um bom resultado quanto à localização e rastreamento dos mesmos.

#####

Trabalhos futuros

Atualmente estamos aprimorando a confiança e continuidade dos dados providos pelo sistema TRUE buscando agregar os dados obtidos através de múltiplos sensores Kinect.

Desta forma podemos evitar as descontinuidades ocorridas por oclusões bem como aprimorar a confiança do reconhecimento obtido.