

6 Discussão

6.1 Intensidade de luz emitida pelo LED

O resistor variável utilizado no circuito elétrico (descrito na seção 3.1.3) utilizado para testar a intensidade dos LEDs, juntamente com o fotospectrômetro Ocean Optics, tornou possível a determinação da intensidade de luz emitida pelos LEDs de 5 e 10 mm de diâmetro, em função da corrente elétrica que passava por eles, com bastante precisão. Para um alcance de corrente entre 90 e 170 mA, o LED de 5 mm de diâmetro apresentou uma variação na intensidade em torno de 5%. Com o mesmo alcance de corrente, a variação da intensidade de luz observada no LED de 10 mm diâmetro foi de 25%. Próximo ao ponto de operação do LED de 10 mm de diâmetro, de 140 mA (valor recomendado pelo fabricante), uma variação na corrente elétrica de $\pm 25\%$ provocou uma flutuação na intensidade de luz desse LED de $\pm 18\%$. A intensidade de luz emitida pelo LED de 10 mm de diâmetro foi da ordem de 18 vezes maior do que aquela medida no LED de 5 mm de diâmetro. Outra vantagem em relação ao uso do LED de 10 mm de diâmetro é que o espectro de luz emitido por esse LED é estreito e apresenta pico máximo de luz em torno de 850 nm, enquanto que o pico máximo de intensidade de luz emitida pelo LED de 5 mm de diâmetro se situa em torno de 930 nm. Isso significa que as imagens capturadas pelo LED de 5 mm incorporam uma quantidade menor de luz na região do infravermelho, que corresponde à faixa de comprimento de onda que deve ser capturada pela filmadora Sony DCR-H21.

A filmadora Sony DCR-H21 demonstrou ser mais sensível à luz com comprimento de onda em torno de 850 nm emitida pelos LEDs de 10 mm (3.6). Uma comparação entre os resultados das imagens feitas com um conjunto de 28 LEDs de 5 mm e com um conjunto de 20 LEDs de 10 mm acoplados à placa de acrílico (5.1) revelou que as imagens capturadas pela filmadora com os LEDs de 5 mm de diâmetro possuem baixo contraste e muitos ruídos, praticamente inviabilizando o funcionamento dos algoritmos utilizados no sistema de visão computacional.

Apesar da equação D8 apresentada no anexo D, comprovada pelos resultados obtidos na seção 5.6.3, indicar uma maior intensidade na luz produzida por um toque sobre a tela quando é utilizada uma configuração na qual os LEDs estão inclinados em

relação à lateral do acrílico, o sistema construído neste trabalho utilizou os LEDs perpendiculares à lateral do acrílico por serem mais simples de construir e serem suficientes para produzir resultados satisfatórios.

6.2 Sistema de projeção

O Projetor Sony VPL-EX4 possui um sistema de lentes para projeção de imagens a curtas distâncias. Porém, apesar disto, não foi possível projetar imagens na superfície da tela de acrílico, que possui uma diagonal igual a 42 polegadas, quando o projetor foi posicionado a uma distância de 1 m da placa, que é a distância de operação do sistema *multitouch*. Portanto, um conjunto de espelhos teve que ser usado para contornar essa dificuldade.

O uso de configurações feitas com espelhos não é recomendado, uma vez que qualquer tipo de poeira ou imperfeição existente na superfície dos espelhos provoca imagens desfocadas na tela do sistema *multitouch* (5.2). Contudo, esse efeito foi minimizado utilizando-se a configuração mostrada na Figura 6.1A. Nessa configuração foi possível projetar imagens refletidas na direção do projetor e, conseqüentemente, minimizar o efeito do desfoque gerado nas imagens formadas sobre a tela.

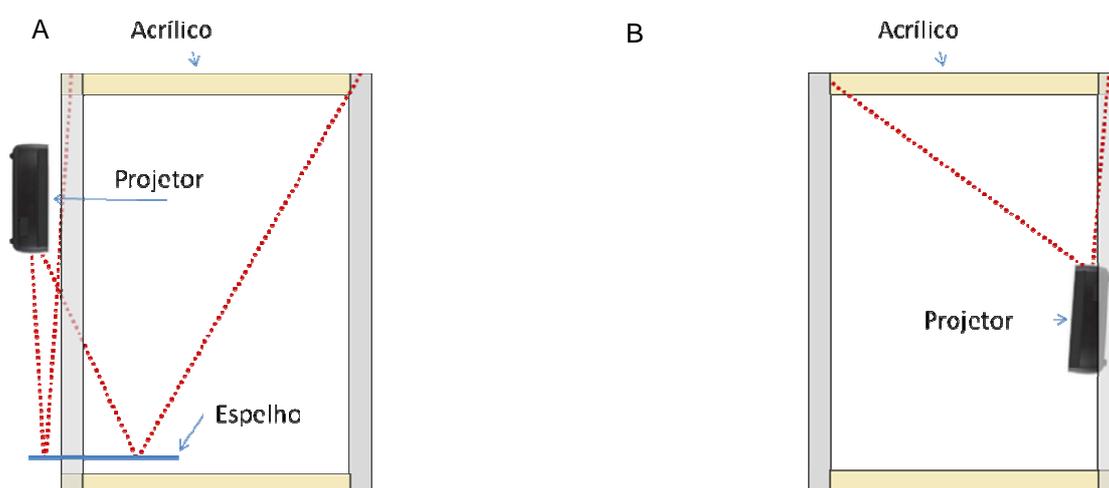


Figura 6.1 – Configurações alternativas do sistema de projeção. Usando um espelho perpendicular ao projetor (A). Usando um projetor com distância de projeção menor (B).

Existem modelos de projetores que permitem o acoplamento de diferentes lentes e, conseqüentemente, a escolha de uma delas para projetar as imagens feitas nesse trabalho a uma distância mais curta. Todavia, devido ao elevado custo dessas lentes, neste estudo foi utilizado um projetor mais simples.

Resultados melhores teriam sido obtidos se os experimentos tivessem sido feitos com um projetor apropriado, por exemplo um BenQ modelo MP771[6]. Isso teria permitido a projeção de imagens de 42 polegadas exigida pelo sistema *multitouch* desenvolvido a uma distância de apenas 60 centímetros da tela de projeção. Um exemplo de configuração utilizando um projetor desse tipo é ilustrado na Figura 6.1B. Nessa configuração, a distância mais curta de projeção possibilita a montagem do sistema de projeção com o projetor posicionado mais próximo da tela ou com o uso de uma tela de até 70 polegadas de diagonal. Esse tamanho de tela é equivalente a uma área com 142 cm de largura e 107 cm de altura.

6.3 Calibração da câmera

A calibração da câmera utilizando uma matriz de homografia demonstrou ser satisfatória. O desvio médio entre as coordenadas nas quais um dedo foi pressionado contra a tela e as coordenadas do mesmo ponto identificado pelo sistema foi de 2,4 mm. Esse valor é considerado satisfatório, uma vez que é menor do que a área de contato do dedo na tela.

Seguindo as recomendações sugeridas por Medeiros [28], o processo de calibração de uma câmera deve ser feito com um número elevado de pontos de calibração. Dessa forma, foi testada a calibração da câmera com quatro e com vinte pontos de calibração. O erro da homografia calculada com os quatro pontos extremos da grade de calibração foi de 2,7 cm, cerca de 10 vezes maior do que aquele obtido na calibração feita com 20 pontos (2,4 mm).

Os resultados apresentados na Tabela 5.2 mostram que o erro de calibração da câmera é maior nos pontos mais afastados do centro da imagem. Isso ocorreu devido ao fenômeno da distorção radial existente na imagem capturada pela filmadora [16].

A Figura 6.2 mostra o efeito da distorção radial. As linhas retas se tornam curvas ao passarem pelo sistema de lentes da câmera. Como pode ser visto na Figura

6.2A, a curvatura das linhas da grade capturadas pela câmera sob a moldura verde externa que aparece na imagem é mais acentuada quando comparada às linhas da moldura menor, desenhadas no centro da tela. A mesma imagem corrigida do efeito da distorção radial é mostrada na Figura 6.2B. Nessa imagem as molduras se alinham à grade que aparece na tela.

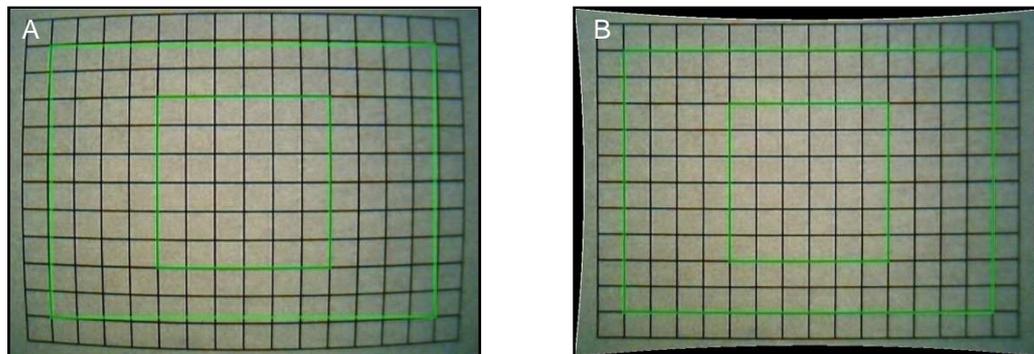


Figura 6.2 – Distorção radial. Imagem original (A). Imagem com correção da distorção radial (B).

O efeito de distorção radial mostrado na Figura 6.2A, no cálculo da matriz de homografia, é minimizado quando um número maior de pontos é utilizado para a calibração da câmera. Contudo, como a homografia é baseada em um modelo linear, esse efeito não pode ser corrigido, apenas minimizado.

A distorção radial depende do tamanho da abertura das lentes angulares. Lentes maiores acarretam distorções radiais mais pronunciadas. Portanto, se em um experimento forem utilizadas câmeras com lente de grande abertura, a distorção deve ser corrigida antes do cálculo da homografia para melhorar os resultados da calibração.

Nos experimentos realizados neste trabalho com a filmadora Sony DCR-H21, o efeito da distorção radial foi mínimo. Dessa forma, não foi necessária a aplicação de uma correção para esse efeito nas imagens obtidas com tal câmara.

Outro fenômeno observado foi que devido ao tempo decorrido entre da captura das imagens pela câmera, o processamento das imagens pelo programa e a posterior projeção dos resultados na tela, existe uma latência de aproximadamente 200 milissegundos entre o ponto real onde o usuário interage com a tela, e onde o sistema considera que o usuário está interagindo. No entanto isto pode ser observado em outros sistemas [22], [14] e é inerente ao uso de visão computacional.

6.4 Algoritmo de rastreamento dos toques

6.4.1 Eliminação da luz de fundo

A abordagem utilizada para filtrar a luz de fundo das imagens capturadas pela filmadora, apesar de simplista, apresentou bons resultados, mesmo em casos em que a imagem capturada possuía muito ruído e regiões mais claras decorrentes da luminosidade proveniente do meio ambiente, que interfeririam na resposta do sistema *multitouch* em estudo.

No início da execução do programa de reconhecimento, o parâmetro α , descrito na seção 4.2.1.3, que controla a taxa de atualização da imagem de fundo, assume um valor igual a 0,066 para os primeiros 60 quadros capturados pela filmadora. Esse valor foi utilizado para obter uma adaptação rápida à imagem de fundo.

Logo após o início da execução do programa de reconhecimento, o valor de α foi reduzido para 0,0025. Isso foi feito para permitir que as flutuações na intensidade da luz ambiente fossem incorporadas à imagem de fundo, que é atualizada a cada novo quadro capturado pela câmera. Com valor de α igual a zero, o modelo de fundo definido durante a inicialização do programa não é atualizada.

Quanto maior o valor de α utilizado no decorrer da execução do programa, maior será a taxa de atualização da luminosidade de fundo nas imagens analisadas. Nesse processo de atualização constante da imagem de fundo, uma imagem estacionária, por exemplo a de um dedo pressionando sobre um ponto da tela, é incorporada à imagem de fundo. Assim, para cada quadro capturado pela câmera, a luminosidade gerada pelo contato do dedo sobre a tela se torna mais intensa na imagem de fundo (fenômeno descrito na Equação 4.14). Como consequência, o valor da luminosidade da região onde ocorre o contato sobre a tela, calculada após a subtração de fundo, que é definida na Equação 4.12, diminui. Após um determinado intervalo de tempo, o valor do sinal se torna baixo demais e é interpretado pelo sistema como sendo o próprio fundo no processo de binarização descrito na seção 4.3.1. Quando foi atribuído o valor 0,0025 para a constante α , o programa demonstrou um melhor desempenho, justificado pela adaptação lenta da imagem de fundo nas imagens em questão. Esse valor da constante α evitou que mudanças na intensidade de luz do ambiente interferissem no sistema *multitouch* e permitiu que um dedo

pudesse ser mantido em uma posição fixa da tela por um período de tempo mais longo antes de ser interpretado pelo programa como sendo parte do fundo na etapa de binarização.

6.4.2 Busca de regiões conexas

A otimização que foi feita na sub-rotina de busca de regiões conexas (Anexo B) permitiu que o valor do filtro da mediana pudesse ser calculado apenas durante a busca recursiva. Dessa forma, grande parte dos pixels “pretos” nas imagens binárias não teve que ser filtrada pela mediana. Adicionalmente, essa otimização tornou possível que a etapa de descarte de regiões, feita com base no número inválido de pixels, também pudesse ser incorporada à sub-rotina de busca. Isso evitou tanto o consumo de memória do computador empregada para armazenar regiões que seriam descartadas em uma etapa seguinte de descarte quanto o custo computacional extra para eliminar as regiões inválidas da lista de regiões encontradas.

6.4.2.1 Binarização da imagem

Quando os LEDs de 5 mm de diâmetro foram utilizados como fonte de luz na tela, a intensidade de luz capturada pela câmera foi demasiadamente fraca. Isso resultou em um baixo contraste entre os pixels relacionados aos dedos sobre a tela e os pixels de fundo. A Figura 6.3A mostra uma imagem do contato de cinco dedos pressionando a tela do sistema *multitouch*. O histograma dessa imagem é mostrado na Figura 6.3B. Nesta imagem os pixels relativos ao contato dos dedos com a tela são demarcados pela faixa amarela que vai até 43% do valor máximo do histograma.

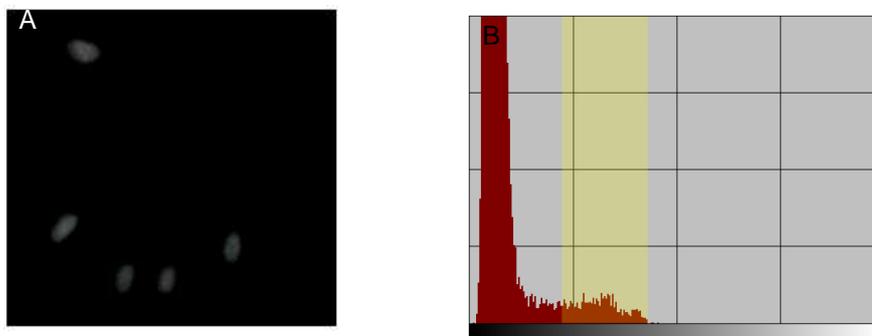


Figura 6.3 – Histograma de uma imagem capturada usando LEDs de 5 mm. Imagem capturada (A). Histograma (B).

O emprego dos LEDs de 10 mm como fonte de luz da tela do sistema *multitouch* melhorou consideravelmente o resultado da binarização. Com esses LEDs, as imagens capturadas pela câmera apresentaram menos ruídos e um maior contraste entre os pixels relativos à imagem dos dedos sobre a tela e o restante da imagem de fundo (Figura 6.4A). A distribuição das intensidades de luz da imagem mostrada na Figura 6.4A é apresentada no histograma da Figura 6.4B. Os pixels relativos às imagens das extremidades dos dedos ocupam uma faixa que vai até 88% do valor máximo do histograma. Os pixels relativos ao contato dos dedos com a tela são demarcados pela faixa amarela.

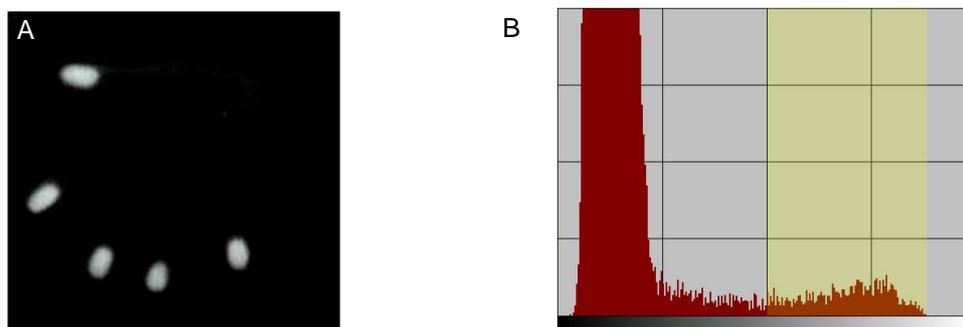


Figura 6.4 – Imagem capturada pela filmadora usando LEDs de 10 mm (A) e respectivo histograma (B).

Com os LEDs de 5 mm de diâmetro, um valor de corte elevado empregado na binarização da imagem, dado pelo parâmetro β (4.3.1), resultavam na perda de pixels nas regiões formadas por um dedo pressionado sobre a tela. Além disso, devido ao melhor contraste observado entre os pixels de fundo e as regiões relativas aos dos dedos sobre a tela iluminada com os LEDs de 10 mm (Figura 6.4A), foi possível atribuir um valor numérico mais elevado para a constante β . Isso evitou que ruídos de fundo fizessem parte da imagem binária sem causar perda dos pixels das regiões correspondentes ao contato dos dedos na tela do sistema *multitouch*.

6.4.2.2 Filtro da mediana

Apesar da melhoria na qualidade das imagens capturadas pela filmadora quando foram utilizados os LEDs de 10 mm de diâmetro, o emprego do filtro da mediana demonstrou ser bastante eficaz na eliminação de seqüências de pixels que ligavam as regiões das imagens dos dedos que tocavam a tela do sistema *multitouch* quando os dedos se encontravam próximos um do outro. O uso desse filtro tornou possível ao algoritmo de busca de regiões conexas identificar e separar as regiões ou

imagens relativas aos toques dos dedos na tela e a imagem de fundo, evitando, dessa forma, o agrupamento delas e a conseqüente interpretação da existência de uma única região (5.4.2).

6.4.2.3 Imagens entrelaçadas

A filmadora Sony DCR-H21 possui uma taxa de captura de 29,97 imagens por segundo. Cada imagem, formada por 720 pixels de largura e 480 pixels de altura, é obtida por duas varreduras na imagem. Durante a primeira varredura, apenas as linhas horizontais pares são atualizadas na imagem. Na segunda varredura, as linhas horizontais ímpares são atualizadas. Esse tipo imagem, que é conhecido como “imagem entrelaçada”, é comumente utilizado nas imagens de televisão.

A Figura 6.5A mostra uma ampliação de uma imagem de quatro dedos se movimentando sobre a tela do sistema *multitouch* capturada por essa filmadora. Na Figura 6.5B podem ser vistos, em amarelo, os pixels obtidos após o processo de binarização da imagem A. As linhas horizontais vistas na imagem B são causadas pela alteração da posição dos dedos sobre a tela, ocorrida em um intervalo de tempo de aproximadamente 17 milissegundos decorridos entre a primeira e a segunda varredura efetuadas pela câmera para formar a imagem entrelaçada. A partir de certa velocidade do movimento do dedo sobre a tela, a imagem binarizada (B) passa a ser formada por um conjunto de linhas horizontais desconexas que inviabilizam a operacionalidade da sub-rotina de busca por regiões conexas no processamento da imagem. Isso ocorre porque cada linha horizontal da imagem é identificada com uma região distinta e é descartada por conter um número baixo de pixels.

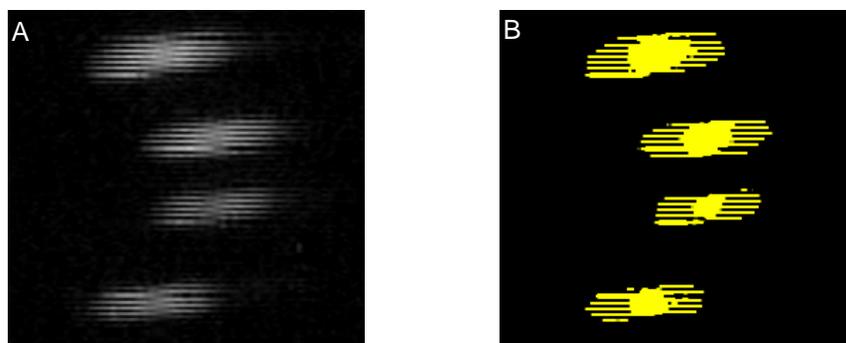


Figura 6.5 – Efeito do entrelaçamento da imagem. Imagem capturada pela filmadora (A). Imagem da filmadora binária (B).

A Figura 6.6 mostra o efeito do desentrelaçamento da imagem capturada pela filmadora. A imagem A contém apenas as linhas horizontais pares obtidas pela primeira varredura da imagem feita pela filmadora. A imagem B contém as linhas horizontais ímpares que foram obtidas na segunda varredura da imagem. Tanto em A como em B, as regiões ou imagens formadas pelo contato dos dedos sobre a tela são claramente definidas, o que permite o reconhecimento dessas imagens pela sub-rotina de busca por regiões conexas. A imagem C representa a soma das imagens A e B, na qual aparece o deslocamento dos dedos durante o intervalo de tempo decorrido entre as duas varreduras feitas pela câmera. No programa desenvolvido neste trabalho, o processamento de imagens entrelaçadas foi feito decompondo cada quadro capturado pela câmera em duas imagens não entrelaçadas, que foram processadas em seqüência. Isso evitou que a sub-rotina de busca por regiões conexas falhasse e permitiu que o programa calculasse a posição de cada região com o dobro da freqüência, o que melhorou o processo de *tracking* de regiões.

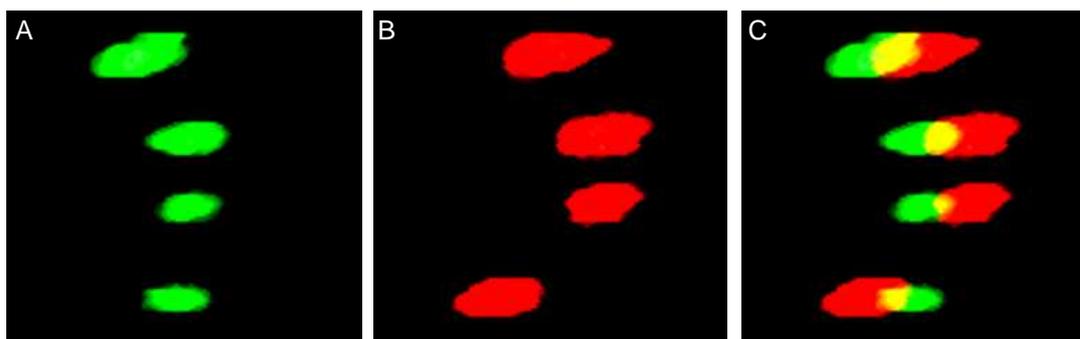


Figura 6.6 – Efeito do desentrelaçamento da imagem. Imagem formada pelas linhas pares (A). Imagem formada pelas linhas ímpares (B). Soma das imagens (C).

6.5

Tracking de regiões

A extrapolação das últimas duas coordenadas de uma região antecessora tornou possível o aumento da velocidade do *tracking* de regiões. Anteriormente, esse procedimento utilizava o cálculo da distância entre as coordenadas de uma dada região encontrada e as da sua antecessora. Contudo, o método utilizando a extrapolação linear demonstrou ser menos eficaz para movimentos curvos dos dedos sobre a tela. Espera-se que o emprego de um modelo não linear de extrapolação, baseado nas medidas das últimas três coordenadas das regiões, melhore os resultados do processo de busca de regiões. Todavia, a solução mais indicada

consiste na aplicação de um filtro de Kalman para prever a próxima coordenada da região em estudo [31].

Para capturar e registrar imagens de um dedo se movimentando rapidamente sobre a tela, a qualidade modesta da filmadora utilizada neste trabalho constituiu um fator bastante limitante no processo de *tracking*. O reconhecimento de uma região na imagem capturada pela filmadora de um dedo se movimentando rapidamente sobre a tela falha antes do teste de correspondência feito com o método da extrapolação. Dessa forma, não foi necessário o uso de um modelo de predição mais avançado.

6.5.1

Broadcast de regiões

O uso da API de mensagens do sistema operacional Windows para a troca de informações entre o programa de reconhecimento e os demais programas feitos para *multitouch* apresentou algumas vantagens. Ao contrário da abordagem feita mediante o uso do protocolo TUIO [9], que é empregado em vários sistemas *multitouch* disponíveis abertamente na Internet (por exemplo, TouchLib [13], OpenTouch [24] e ReactiVision [19]), a abordagem utilizada neste trabalho não exigiu as etapas de conexão entre o programa de reconhecimento e os aplicativos *multitouch*. Os aplicativos que usam TUIO devem tratar as possíveis perdas de pacotes de dados UDP enviados pelo protocolo (empregados nos sistemas acima mencionados). O sistema proposto simplificou consideravelmente a programação dos aplicativos *multitouch*, uma vez que há garantia de entrega das mensagens. Outra vantagem do emprego da API de mensagens do sistema operacional Windows está relacionada ao envio e recebimento das mensagens. As mensagens enviadas pelo programa de reconhecimento puderam ser transmitidas para todos os aplicativos do sistema operacional. Isso tornou possível o uso simultâneo de múltiplos aplicativos *multitouch*.

No programa computacional *multitouch* desenvolvido neste trabalho, cada ponto de contato da tela identificado pelo programa de reconhecimento foi mapeado para o mesmo sistema de coordenadas que é utilizado pelo projetor antes de ser enviado para os demais aplicativos. Esse sistema de coordenadas é o mesmo que aquele utilizado pelo sistema operacional Windows para definir o posicionamento dos aplicativos na imagem projetada. Adicionalmente, as coordenadas recebidas na mensagem do programa de reconhecimento são convertidas pelo aplicativo *multitouch* para o seu próprio sistema de coordenadas, determinado pela região da tela ocupada pelo aplicativo. Quando as mensagens não são identificadas como sendo pertinentes

à área do aplicativo na tela, elas são ignoradas. Isso permitiu que cada aplicativo *multitouch* ativo fosse influenciado apenas por toques que ocorressem dentro de uma região delimitada por eles na imagem projetada na tela.

6.5.2 Integração com o sistema operacional Windows

O sistema operacional Windows permite que um dado aplicativo utilize a função programada em C++ *SendInput()* para emular o movimento do mouse ou o uso do teclado de um computador [27]. Neste trabalho, a função *SendInput()* foi definida como:

```
UINT SendInput(  
    UINT nInputs,  
  
    LPINPUT pInputs,  
  
    int cbSize  
  
);
```

O emprego da função *SendInput*, após o processo de reconhecimento dos pontos sobre a tela, tornou possível o uso das informações dos toques para controlar a posição do mouse no sistema operacional, utilizando-se os valores das coordenadas da maior região presente na tela. Adicionalmente, o evento de clique do botão do mouse foi gerado quando são encontradas duas regiões próximas. A região com maior tempo de vida pode ser utilizada para determinar as coordenadas do cursor.

A digitação de textos na tela do sistema *multitouch* pode ser feita mediante o emprego de um programa *multitouch* que mostre um teclado virtual sobre a tela. Esse teclado virtual, ao receber um toque sobre uma determinada área da tela que corresponda a uma letra do teclado, emula a digitação através do teclado real. Isso permite que o sistema operacional do computador seja controlado exclusivamente pelos toques sobre a tela do sistema *multitouch*, evitando a necessidade do uso do mouse ou teclado.

6.6 Tela do sistema *multitouch*

O processo utilizado para o revestimento do acrílico com a parafina em gel não resultou em uma camada uniforme do gel sobre a tela. Isto foi atribuído à falta de uniformidade da fonte de calor (quatro lâmpadas de infravermelho de 250 W). Contudo, o uso de uma estufa possibilitaria uma distribuição uniforme do gel sobre a superfície do acrílico.

O perfil em “U” de alumínio com espessura interna de 11 mm, onde foram fixados os LEDs de 10 mm de diâmetro, juntamente com a camada de parafina de aproximadamente 2 mm de espessura, possibilitou o acoplamento dos LEDs, sem folga, nas laterais da placa de acrílico.

A homogeneidade do campo de luz sobre a superfície do acrílico foi determinada experimentalmente com a câmera e comparada com os resultados previstos pela Equação 6.1, que determina a profundidade de penetração, d , da onda evanescente formada na interface entre o acrílico e o ar.

$$d = \frac{\lambda}{2\pi \sqrt{n_i^2 \text{Sen}^2(\theta_i) - n_t^2}} \quad (6.1)$$

λ é comprimento de onda, n_i e n_t são os índices de refração dos meios do acrílico e da parafina, respectivamente, e θ_i é o ângulo de incidência do raio entre as interfaces dos meios. Uma explicação detalhada da Equação 6.1 é apresentada no Anexo D.

As medidas apresentadas neste trabalho poderiam ter sido feitas em uma tela de menor dimensão. Contudo, o tamanho da tela utilizada demonstrou ser útil para os propósitos deste estudo e ela poderá vir a ser utilizada em outros experimentos que requirem a manipulação simultânea da tela por vários usuários ou a simulação de objetos grandes, tais como o teclado de um piano virtual.