

Ferramenta de Labeling - Manual do Utilizador

30 de Setembro de 2014

Conteúdo

1	Versão do software	2
2	Introdução	2
3	Instalação e compilação da ferramenta	3
3.1	Obter a ferramenta (SVN)	3
3.2	Dependências e sistemas suportados	3
3.3	Compilar a ferramenta	3
4	Execução da ferramenta	3
4.1	Linha de comando	3
4.2	Ficheiro de video	4
4.3	Ficheiro de labeling	4
4.4	Autosave e backup	4
5	Utilização da ferramenta	5
5.1	Utilização geral	5
5.1.1	Ajuda (tecla h)	5
5.1.2	Teclado	5
5.1.3	Rato	5
5.1.4	Gravar	6
5.1.5	Sair	6
5.1.6	Navegar no vídeo	6
5.1.7	Zoom	7
5.1.8	Tamanho da janela	9
5.1.9	ID dos objectos	9
5.2	Labels	9
5.2.1	Anatomia duma label	9
5.2.2	Tipos de labels	9
5.2.3	Criar labels	11
5.2.4	Apagar labels	11
5.2.5	Modificar labels	11
5.2.6	Finalizar as labels	13
5.3	Automatismos	14
5.3.1	Cópia	14
5.3.2	Cópia com procura	14
5.3.3	Interpolação linear	14

5.3.4	Procura numa sequência	15
5.4	Verificação e confirmação da labels	16

Lista de Figuras

1	Ajuda	5
2	Mapa do teclado	6
3	Informação	7
4	Zoom	8
5	ID's dos objectos	10
6	Tipos de labels.	10
7	Formas de usar o rato.	12
8	Interpolação linear	15
9	Ajuste automático	16

1 Versão do software

Este manual documenta a versão 0.1 da ferramenta (labelingtool v0.1).

2 Introdução

Tanto para treino dos algoritmos de visão como para o teste dos mesmos, é necessário utilizar sequências de imagens que tenham sido previamente anotadas.

A tarefa de anotação (labeling) consiste em marcar um rectângulo em volta de todas os objectos que aparecem nas imagens. Tal operação tem de ser feita manualmente, pois serve como base de verdade (Ground Truth) sobre a qual os algoritmos vão ser treinados e testados. Qualquer erro no processo de etiquetagem afectará negativamente a performance dos algoritmos e comprometerá a fiabilidade das avaliações efectuadas.

Este processo manual é bastante moroso, especialmente se o número de imagens for elevado como normalmente é o caso nas sequências de vídeo. Por esse motivo, esta ferramenta foi desenvolvida tendo em vista tentar aliviar o peso desta tarefa. Isto foi conseguido de várias formas. Em primeiro lugar, criando uma interface que seja leve e rápida e tentando minimizar o número de operações necessárias por cada anotação.

Isto foi conseguido desenvolvendo a ferramenta na linguagem C++ (foram usadas as bibliotecas do OpenCV). Houve o cuidado de tornar a execução de todas as operações o mais rápido possível, de forma a que o operador veja imediatamente o resultado dos comandos que executa e de forma a que possa executar comandos em sequência de forma rápida. As ferramentas manuais existentes anteriormente, programadas em Matlab, eram demasiado morosas.

Esta ferramenta é genérica, podendo-se aplicar a qualquer tipo de imagens para etiquetar qualquer tipo de objectos. Tem a limitação de apenas suportar labels rectangulares.

3 Instalação e compilação da ferramenta

3.1 Obter a ferramenta (SVN)

A ferramenta está na directória

```
devel/tools/labeling_tool
```

no SVN do projecto SEAGULL nos servidores do ISR. Para descarregar a ferramenta, usar o comando:

```
svn checkout svn://svn.isr.ist.utl.pt/seagull/devel/tools/labeling_tool
```

3.2 Dependências e sistemas suportados

Esta ferramenta depende das bibliotecas do OpenCV, portanto é necessário instalá-las antes de compilar.

A ferramenta foi desenvolvida em linux, portanto compila e funciona bem neste tipo de sistemas.

Foi também compilada e testada em Mac OS X. Esta versão é utilizável mas apresenta alguns problemas, especialmente no que diz respeito ao tamanho da janela. Portanto a funcionalidade descrita na subsecção 5.1.8 fica afectada em sistemas Mac OS X.

3.3 Compilar a ferramenta

Para compilar a ferramenta, execute os seguintes comandos:

```
cd labeling_tool  
mkdir build  
cd build  
cmake ..  
make
```

4 Execução da ferramenta

4.1 Linha de comando

A ferramenta é executada dentro da directória build (ver subsecção 3.3) usando o comando

```
./labeling_tool [options] <video_file>
```

Para obter ajuda usar

```
./labeling_tool -h
```

4.2 Ficheiro de video

O ficheiro de vídeo pode ser qualquer ficheiro num formato legível pelo OpenCV (exemplos: .mpg, .avi, .mkv, etc.). Indicar o ficheiro no comando. Exemplo:

```
./labeling_tool lanchaArgos_clip1.avi
```

Se o ficheiro não estiver na pasta actual, indicar também o caminho:

```
./labeling_tool videos/lanchaArgos_clip1.avi
```

4.3 Ficheiro de labeling

Por defeito, o nome do ficheiro onde as labels são guardadas é igual ao nome do ficheiro de vídeo (incluindo o caminho) mas com a extensão mudada para .gt.txt. Por exemplo, o comando

```
./labeling_tool videos/lanchaArgos_clip1.avi
```

cria o ficheiro de labeling

```
videos/lanchaArgos_clip1.gt.txt
```

Em alternativa pode-se usar a opção -d para especificar um nome de ficheiro diferente:

```
./labeling_tool -d labels.gt.txt lanchaArgos_clip1.avi
```

O formato do ficheiro (labelingtool v0.1) consiste num ficheiro de texto com várias linhas, uma por cada label. O formato de cada linha é o seguinte:

```
<frame number> <x> <y> <width> <height> <object id> <1(temporary)/0(final)>
```

O formato do ficheiro também está descrito na ajuda da ferramenta.

```
./labeling_tool -h
```

4.4 Autosave e backup

A cada 30 segundos é gravado um ficheiro com a extensão .autosave (exemplo: labels.gt.txt.autosave) com todas as labels e com todas as alterações feitas até esse ponto. Esse ficheiro é apagado quando o programa sai normalmente.

É ainda feita uma cópia de backup do ficheiro original antes de se gravar o ficheiro novo. Essa cópia tem o nome igual ao original mas com um til (~) adicionado no fim (exemplo: labels.gt.txt~).

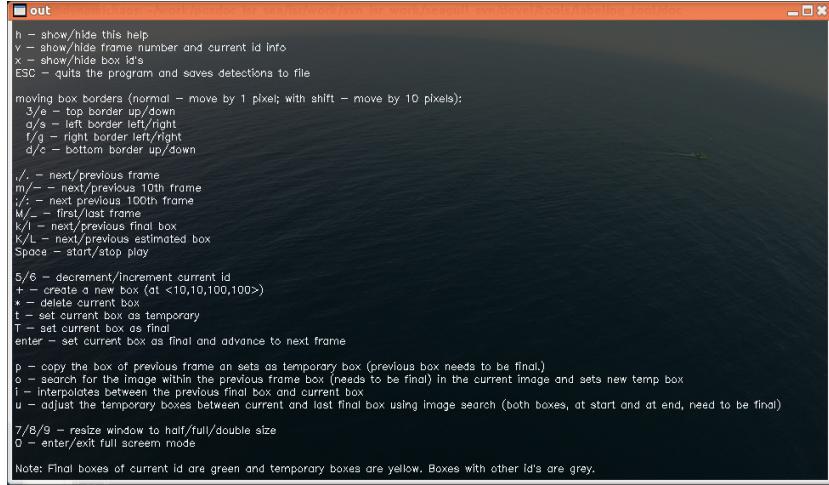


Figura 1: Ajuda - carregar em “h”.

5 Utilização da ferramenta

5.1 Utilização geral

5.1.1 Ajuda (tecla h)

Está sempre disponível em qualquer instante um ecrã de ajuda (Figura 1) onde estão identificadas todas as teclas (ou combinações de teclas) para todos os comandos possíveis assim como uma breve descrição dos mesmos. Carregue na tecla ‘h’ para chamar esse ecrã de ajuda, carregue novamente para sair do ecrã.

A ferramenta continua a trabalhar normalmente mesmo com a ajuda visível. Além disso, o ecrã de ajuda é translúcido. Isto permite que se continue a poder executar outros comandos e a fazer ajustes com o rato.

5.1.2 Teclado

Os comandos são na sua quase totalidade efectuados através do teclado (ver Figura 2), existindo teclas únicas ou combinações simples de teclas para realizar todas as operações necessárias. Foram ainda levadas em conta questões ergonómicas, tentando-se reduzir o número de movimentos necessários não só por cada comando mas também nas sequências de comandos mais utilizadas.

O mapa da Figura 2 está disponível junto com o código fonte na pasta images

`labeling_tool/images/help.svg`

e pode ser impresso para referência.

5.1.3 Rato

É ainda possível usar o rato do computador, mas essa possibilidade é mais apta para movimentações grandes das labels no ecrã. Esse tipo de movimentação é rara, sendo preferível manter ambas as mãos no teclado ao invés de desperdiçar tempo a mudar a mão entre o teclado e o rato.

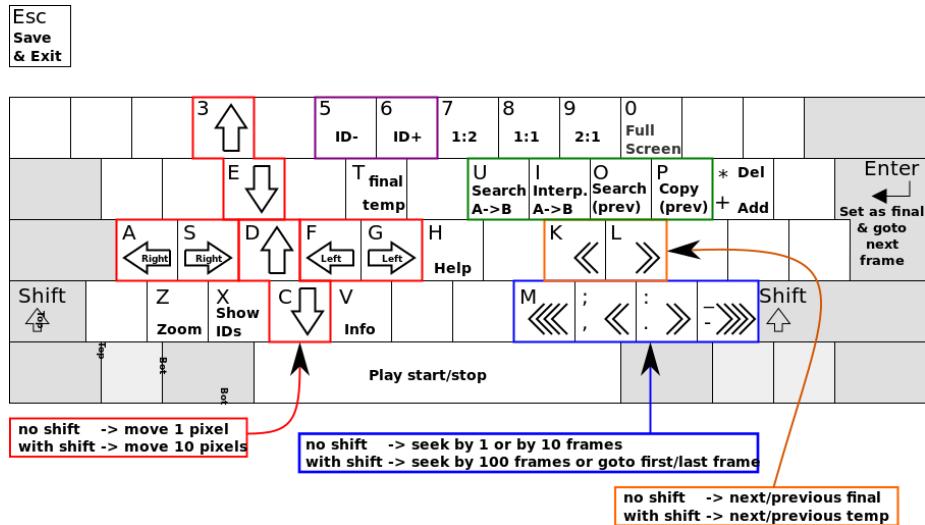


Figura 2: Mapa do teclado mostrando os comandos disponíveis na ferramenta de labeling (mapa disponível em labeling_tool/images/help.svg).

O rato pode ser usado para mover ou para modificar a label. Para tal, carrega-se no botão direito, arrasta-se o rato e larga-se o botão na nova posição desejada. O ponto na imagem sobre o qual o cursor do rato está quando se carrega no botão define a operação realizada pelo rato. Para mais informação ver a subsecção 5.2.5.

5.1.4 Gravar

Contrariamente ao que é habitual, não existe nenhum comando para gravar o ficheiro de labels. O ficheiro é gravado automaticamente quando se sai do programa (carregar na tecla Escape).¹

A cada 30 segundos é gravado um ficheiro com a extensão .autosave com todas as labels e com todas as alterações feitas até esse ponto. Esse ficheiro é apagado quando o programa sai normalmente.

5.1.5 Sair

Carregar na tecla Escape. O ficheiro de labels é gravado automaticamente.

5.1.6 Navegar no vídeo

Existem várias teclas que permitem avançar e retroceder no tempo do vídeo. Isto pode ser feito:

- tocar e parar o vídeo – tecla de espaços
- avançar/retroceder de frame a frame – teclas “,” e “.”

¹**Dica:** para sair do programa sem gravar, ir para o terminal onde o programa foi lançado e aí carregar em CTRL+”c”. Desta forma o programa é imediatamente morto sem ter a oportunidade de gravar o ficheiro.

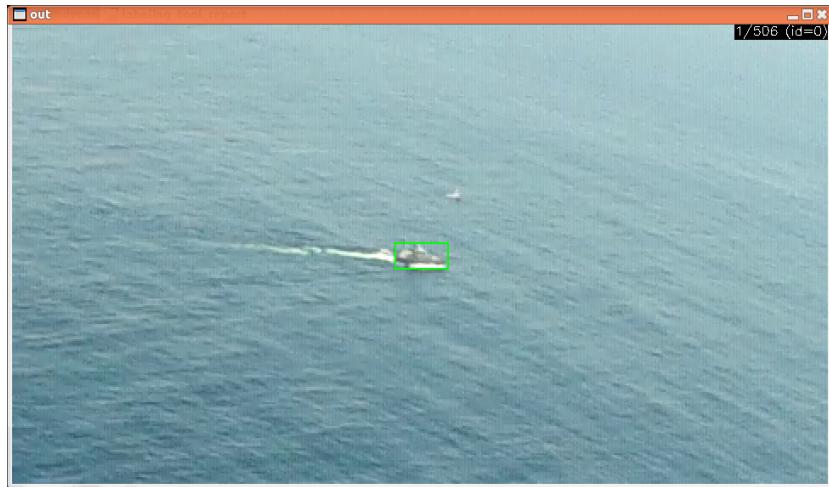


Figura 3: Informação - ao carregar na tecla “v” o numero da frame e ID do objecto são apresentados no canto superior direito.

- avançar/retroceder de 10 em 10 frames – teclas “m” e “-”
- avançar/retroceder de 100 em 100 frames – teclas “;” e “:” (ou seja, shift + “;” e shift + “.”)
- ir para a primeiro frame do vídeo – tecla “M” (ou seja, shift + “m”)
- ir para a última frame do vídeo – tecla “_” (ou seja, shift + “-”)

É ainda possível procurar a labels já feitas. Pode-se ir para a frame anterior ou para a próxima frame que:

- contenha uma label temporária - teclas shift + “k” e shift + “l”
- contenha uma label final - teclas “k” e “l”

Para saber qual é a frame actual, carregue na tecla “v” e o número da frame aparecerá no canto superior direito da imagem (Figura 3).

5.1.7 Zoom

Para se conseguir uma maior precisão, é útil ver em pormenor a zona da imagem onde se está a colocar ou a ajustar uma label. Existe para isso a tecla “z” que sendo pressionada várias vezes muda o zoom da imagem ciclicamente entre 1 (sem zoom), 2 e 4 vezes (Figura 4). Todos os comandos funcionam independentemente do zoom utilizado.

Este zoom é centrado automaticamente de forma a que a label marcada fique centrada no ecrã. Qualquer ajuste da label provoca uma nova centragem.

Nota: Este comando de zoom não altera o tamanho da janela. Para isso ver a subsecção 5.1.8.

Aviso1: Devido à auto-centragem, pode se tornar pouco natural fazer ajustes da label usando o rato quando o zoom está activo.



(a) 1x



(b) 2x



(c) 4x

Figura 4: Zoom - carregar na tecla “z” muda o zoom ciclicamente:
 $1x \rightarrow 2x \rightarrow 4x (\rightarrow 1x \rightarrow \text{etc.})$.

Aviso2: Ao navegar no vídeo com o zoom activo, a imagem pode saltar de forma imprevisível. Este comportamento é normal e acontece quando a frame actual não contém nenhuma label, tornando impossível a auto-centragem nessa frame. Solução: escolher o zoom 1x ou ignorar os saltos da imagem.

5.1.8 Tamanho da janela

(Ver a subsecção 3.2 sobre as limitações em Mac OS X.)

As teclas “7”, “8” e “9” permitem definir o tamanho da janela para metade, igual ou o dobro do tamanho original do vídeo. Este tamanho é independente do zoom escolhido (ver subsecção 5.1.7), sendo sempre relativo ao tamanho do vídeo com o zoom igual a 1x.

É ainda possível colocar a janela a ocupar todo o ecrã (full-screen) usando a tecla “0”. Neste modo maximiza-se a utilização do ecrã.

Nota: para obter uma representação do vídeo no ecrã pixel a pixel, escolher o tamanho igual ao original (tecla “8”) e seleccionar o zoom igual a 1x (subsecção 5.1.7).

Nota: por defeito, a ferramenta é inicializada com uma janela de tamanho igual a metade do tamanho do vídeo. Isto ajuda quando os vídeos são maiores que o ecrã.

Nota: no sistema operativo OS X esta funcionalidade ainda tem problemas (ou seja, não funciona).

5.1.9 ID dos objectos

É possível fazer o labeling de vários objectos na mesma imagem. A cada objecto atribui-se um número de identificação (id). O id actual é escolhido usando as teclas “5” e “6” que o incrementam e decrementam.

A label correspondente ao id actual tem cor verde ou laranja, conforme o tipo (ver subsecção 5.2.2). As labels dos outros id's são mostradas a cor cinzenta.

Carregando na tecla “x”, é mostrado junto das labels qual o id a que cada label corresponde (Figura 5 e Figura 6d). Carregando em “v” é também mostrado no canto superior direito o ID seleccionado.

Todos os comandos afectam apenas e somente a label do id actual. Para criar, modificar ou apagar uma label com outro id que não seja o actual, então é obrigatoriamente necessário usar as teclas “5” e “6” para seleccionar o id correcto.

5.2 Labels

5.2.1 Anatomia duma label

Uma label consiste num rectângulo com quatro fronteiras. O objecto deve ficar dentro da label e a label deve ser a menor possível que consiga conter o objecto completo.

Esta ferramenta permite não só mover a label completa mas também ajustar cada uma das suas fronteiras independentemente.

5.2.2 Tipos de labels

Existem dois tipos de labels (ver Figura 6):

- labels temporárias - cor laranja

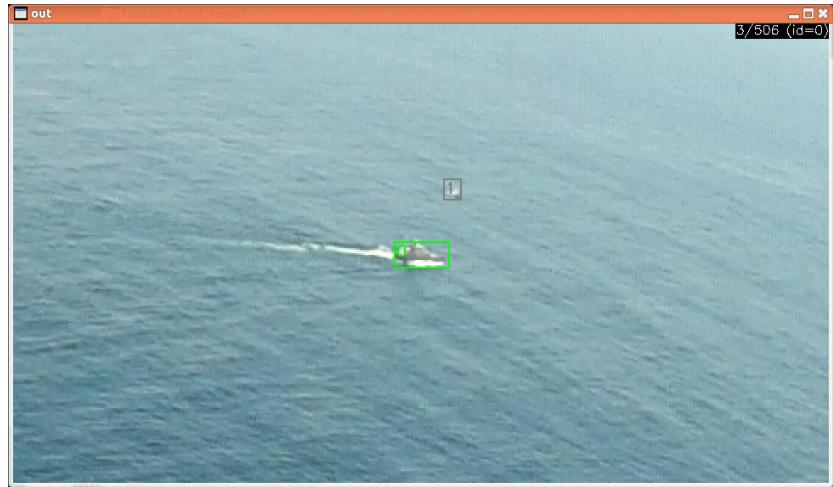


Figura 5: ID's dos objectos – ao carregar na tecla “x” os ID's são mostrados dentro dos rectângulos das labels. Carregando em “v” é mostrado no canto superior direito o ID seleccionado (neste caso, ID=0).

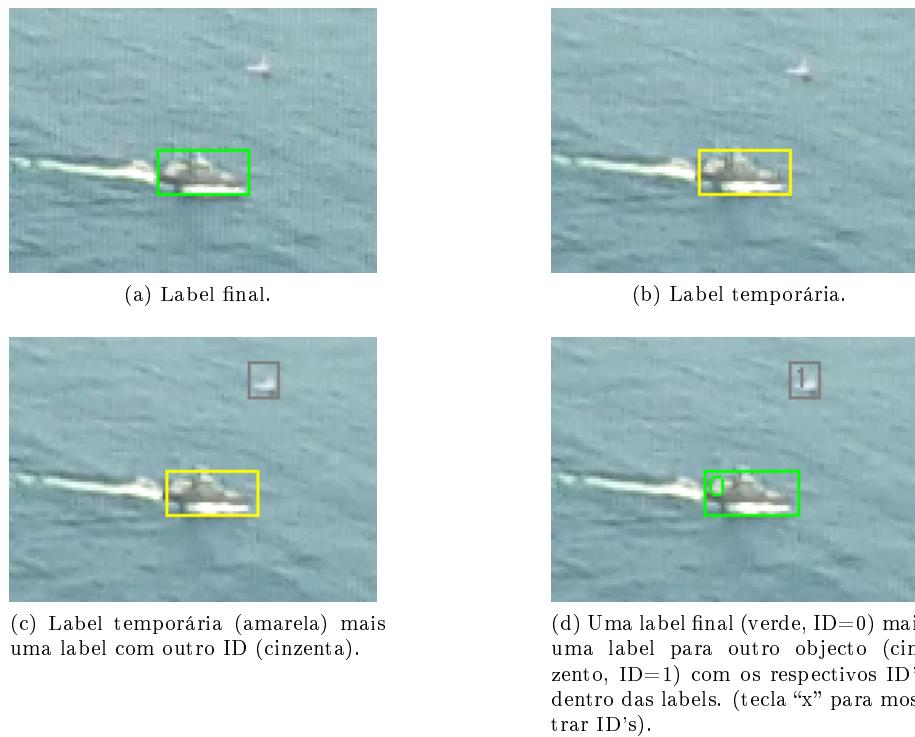


Figura 6: Tipos de labels.

- labels definitivas - cor verde

Sempre que uma label é criada, o seu tipo é temporário. Isto inclui labels criadas manualmente (subsecção 5.2.3) ou automaticamente (secção 5.3).

As labels só se tornarão definitivas quando:

- forem ajustadas pelo operador (subsecção 5.2.5)
- forem aceites pelo operador como sendo definitivas (subsecção 5.2.6)

Nota: existiu a preocupação de minimizar o número de operações necessárias por label. Para tal, vários dos comandos da ferramenta de labeling transformam automaticamente as labels temporárias em definitivas sem ser necessário ao operador fazê-lo explicitamente. Cabe ao operador se assegurar que todos os ajustes necessários foram aplicados a todas as labels definitivas.

5.2.3 Criar labels

As labels podem ser criadas de duas formas:

- manualmente - tecla “+”
- automaticamente - (ver subsecção 5.3)

Quando criadas manualmente, a label é colocada por defeito no canto superior esquerdo da imagem. O operador deverá então movê-la para o local certo e ajustar as suas fronteiras, usando para isso o teclado ou o rato.

Dica: provavelmente neste caso (criação manual) será melhor usar em primeiro lugar o rato para fazer um ajuste grosso da label e depois usar o teclado para fazer um ajuste mais fino (usando zoom).

5.2.4 Apagar labels

Carregue nas teclas shift e “+” (ou seja, na tecla “*”) para apagar a label actual.

5.2.5 Modificar labels

Usar o rato O rato pode ser usado para mover ou para modificar a label. Para tal, carrega-se no botão direito, arrasta-se o rato e larga-se o botão na nova posição desejada. O ponto na imagem sobre o qual o cursor do rato está quando se carrega no botão define a operação realizada pelo rato.

Mover a fronteira da label usando o rato Se se carregar sobre uma das fronteiras ou sobre uma das rectas imaginárias que passam sobre as fronteiras, então o rato vai mover a fronteira em questão. Ao colocar o rato sobre uma fronteira ou sobre a respectiva recta imaginária, a cor da fronteira muda para vermelho assinalando que essa será a fronteira seleccionada ao premir o botão do rato.

Apontar o rato para a recta imaginária (fora da fronteira) é útil quando a label é demasiado pequena e se torna difícil acertar com o rato na fronteira propriamente dita.

É ainda possível mover duas fronteiras em simultâneo apontando o rato para o canto da label onde as duas fronteiras se cruzam.

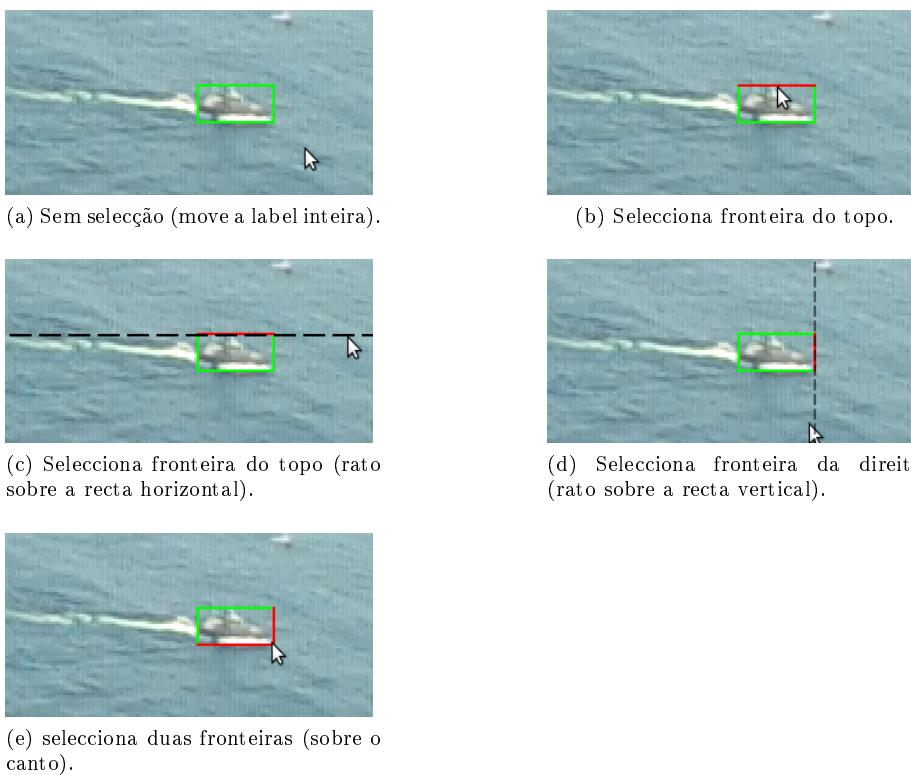


Figura 7: Formas de usar o rato.

Mover a label usando o rato Carregando em qualquer outro ponto que não seleccione nenhuma fronteira, o rato move a label por inteiro sem alterar as suas dimensões.

Mover a fronteira da label usando o teclado Existem teclas independentes para mover cada uma das fronteiras em cada uma das direcções (ver Figura 2). A disposição dessas teclas está organizada de uma forma lógica tanto no que diz respeito à fronteira a mover como na direcção do movimento.

Por exemplo, ambas as teclas para mover a fronteira do lado esquerdo (“a” e “s”) estão do lado esquerdo relativamente às outras teclas, e a tecla que move essa fronteira para a esquerda é o “a” que está à esquerda da tecla “s” que move a fronteira para a direita. Espera-se que desta forma o operador não precise de sequer olhar para o teclado para ajustar a label.

Existe ainda a possibilidade de mover a fronteira de 10 em 10 pixéis, carregando em simultâneo na tecla SHIFT. Desta forma conseguem-se fazer movimentos mais rápidos carregando menos vezes nas teclas.

Mover a label usando o teclado Não existem teclas específicas para mover a label como um todo. No entanto, pode-se facilmente conseguir esse efeito usando as teclas que movem as fronteiras.

Por exemplo, vamos supor que se pretende deslocar a label para o lado esquerdo. Esta operação é equivalente a deslocar tanto a fronteira esquerda como a direita, ambas para o lado esquerdo. Podemos então usar as teclas “a” e “f”, e se as carregarmos em rápida sucessão, o movimento que resulta é o pretendido. Melhor ainda, na maioria dos casos (dependendo do hardware) podemos carregar nas duas teclas em simultâneo porque o teclado vai automaticamente enviar os dois keypresses um a seguir ao outro (não interessa qual vai primeiro desde que sigam o mesmo numero de keypresses para cada uma das fronteiras).

Esta técnica também funciona carregando em simultâneo na tecla SHIFT, conseguindo-se mover a label de 10 em 10 pontos.

5.2.6 Finalizar as labels

Como já foi referido, as labels são criadas como sendo do tipo temporário. Isto acontece tanto na caso da label ser adicionada manualmente ou automaticamente. Cabe depois ao utilizador verificar todas as labels, ajustá-las conforme necessário e marcá-las como finais.

Para efeitos de optimização do trabalho, tentou-se minimizar o numero de teclas que é preciso carregar por cada label. Nessas sentido, algumas operações, nomeadamente as que movem a label ou as suas fronteiras, automaticamente mudam o tipo da label de temporária para final. A lógica assumida é que se o operador ajustou a label então é porque já a colocou, ou está no processo de a colocar, na sua posição final.

Para os casos em que não é necessário nenhum ajuste, pode-se carregar na tecla “t” para tornar a label final.

Pode-se ainda carregar em SHIFT+”t” para tornar a label temporária novamente.

5.3 Automatismos

Uma forma de aliviar a tarefa de labeling, foi criando automatismos que, embora não pretendam substituir a anotação manual, pretendem facilitar a mesma.

Foram criados quatro métodos automáticos, dos quais dois operam imagem a imagem enquanto que os outros dois automatizam a anotação de um número elevado de imagens num único passo.

Estes métodos são os que se seguem.

5.3.1 Cópia

Carregar na tecla “p”.

Este comando copia a anotação do frame de vídeo anterior. Essa frame anterior precisa de existir e ser final, caso contrário o comando não faz nada.

Pressupõe-se que os barcos se deslocam pouco entre imagens consecutivas, logo é preferível copiar a anotação da imagem anterior e fazer pequenos ajustes (para os quais existem teclas dedicadas) do que criar uma nova anotação desde o início para cada imagem.

5.3.2 Cópia com procura

Carregar na tecla “o”.

Este comando é idêntico ao anterior, mas depois de fazer a cópia faz uma procura da imagem dentro do rectângulo definido pela label do frame anterior. Essa procura faz-se no frame actual e numa vizinhança à volta da label original. A procura é feita com base na minimização do erro quadrático médio entre as duas imagens.

Na maioria dos casos, este método produz uma anotação muito boa, requerendo ocasionalmente alguns pequenos ajustes.

5.3.3 Interpolação linear

Neste método são marcadas duas anotações em duas imagens não consecutivas, sendo depois feita uma interpolação linear para estimar as anotações das imagens intermédias. Ver o exemplo da Figura 8.

Para este método são necessários os seguintes passos:

1. marcar e finalizar a label na primeira frame do intervalo
2. marcar e finalizar a label na última frame do intervalo
3. ir para (seek) a última frame do intervalo
4. carregar na tecla “i”

Ao carregar na tecla “i”, o método anda para trás no vídeo até encontrar uma label do tipo final. Depois faz uma interpolação linear entre essa label que encontrou e a label actual, criando labels temporárias novas ou alterando as labels temporárias já existentes.

Este método aplica-se aos casos onde o objecto descreve movimentos lineares. Muitos dos movimentos encontrados frequentemente ou são deste tipo linear ou podem ser aproximados por troços lineares. Este método em conjunto com o seguinte (subsecção 5.3.4) permite na maioria dos casos obter aproximações

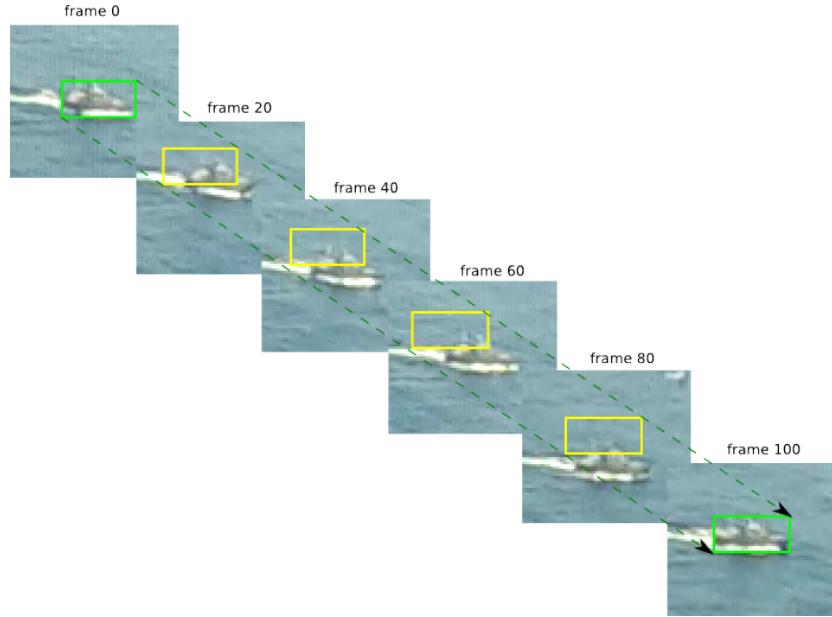


Figura 8: Exemplo de anotações obtidas automaticamente usando o método da interpolação linear entre os frames 0 e 100. As anotações extremas (frames 0 e 100, a cor verde) são manuais.

muito boas, minimizando o número de ajustes manuais necessários após a sua aplicação.

5.3.4 Procura numa sequência

Este método ajusta labels temporárias (por exemplo, criadas usando o método anterior – subsecção 5.3.3) predefinidas num intervalo entre duas frames com label finais já definidas. Ver exemplo na Figura 9.

Para este método são necessários os seguintes passos:

1. marcar como finais as duas labels nos extremos do intervalo de frames
2. criar as labels temporárias dentro do intervalo
3. ir para (seek) a última frame (frame final do fim do intervalo)
4. carregar na tecla “u”

Ao carregar na tecla “u”, o método anda para trás no vídeo até encontrar uma label do tipo final, definindo assim o intervalo de frames onde opera. Depois, guarda as imagens dos extremos do intervalo (as que estão dentro das labels finais).

Para cada frame dentro do intervalo, o método procura as duas imagens numa vizinhança à volta da label temporária predefinida nessa frame. A procura faz-se com base no erro quadrático médio, semelhante ao método da subsecção 5.3.2, mas agora são obtidos dois erros diferentes, um por cada imagem. Os erros são pesados com base na distância entre e frame actual e as frames das

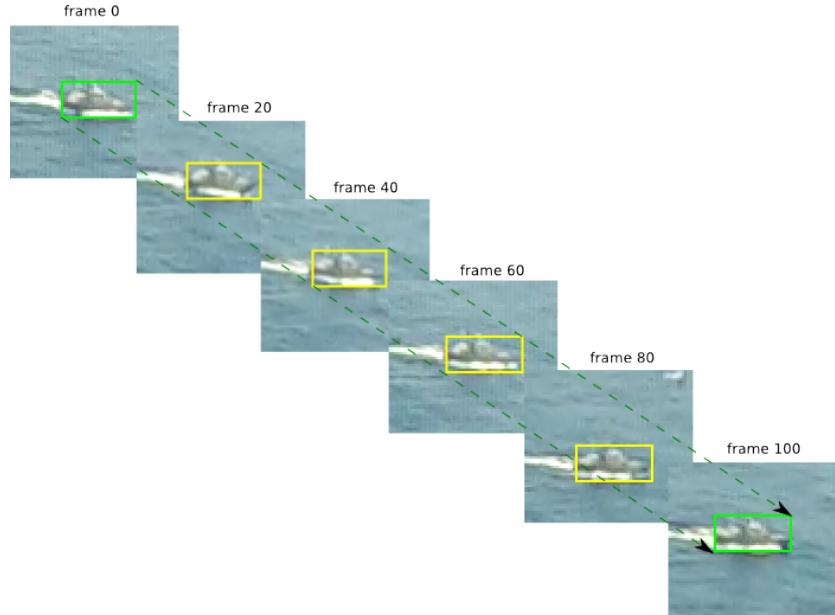


Figura 9: Anotações obtidas depois de ajustadas automaticamente (relativamente à figura anterior).

duas imagens e por fim o erro mais baixo é escolhido. A posição encontrada que corresponde à imagem com menos erro é usada como sendo a nova label temporária para essa frame.

O peso com base na distância temporal parte do pressuposto que a imagem do objecto será mais semelhante quanto mais perto temporalmente forem os dois frames comparados. Usando as duas imagens dos extremos do intervalo e usando este peso, permite ao método funcionar em intervalos de tempo mais alargados mesmo que a imagem do objecto varie nesse intervalo (por exemplo: o objecto pode rodar, mudar de tamanho devido à perspectiva, etc.).

Este método adequa-se bem a ser utilizado com o anterior (subsecção 5.3.3) pois compensa os desvios do objecto em relação à trajectória linear. Frequentemente a trajectória do objecto na imagem não é exactamente linear. Isto pode acontecer por a trajectória do objecto não o ser de facto (por exemplo, trajectória circular) ou pode acontecer por outros motivos externos ao objecto como, por exemplo, vibrações ou outras movimentações da câmara que captura as imagens.

Este método foi aplicado com bastante sucesso em sequências com até 500 frames.

5.4 Verificação e confirmação da labels

Note-se ainda que estes métodos não dispensam por parte do operador a verificação manual (imagem a imagem) das anotações estimadas automaticamente. Aliás, a ferramenta marca-as como anotações temporárias, forçando o operador a torná-las finais manualmente. No entanto, os métodos disponíveis na ferramenta estimam correctamente um grande número das anotações, restando ao

operador fazer pequenas correções ocasionais e confirmar as anotações corretas.