

MG *tool* - Manual do utilizador

1 Introdução

Este documento tem como objetivo tornar o utilizador mais familiarizado com a aplicação. Serão apresentadas as principais características da ferramenta para que o utilizador retire maior partido das suas potencialidades.

A aplicação foi desenvolvida num computador com o sistema operativo 32bits *Windows* 7. Utilizando tecnologia *Java*, através do IDE *Eclipse Helios*. Para permitir visualizar as sequências de vídeo utilizou-se a biblioteca *Java Media Framework* (JMF). Esta biblioteca permite trabalhar com vários formatos de vídeos, que são descritos em <http://www.oracle.com/technetwork/java/javase/formats-138492.html>. Para testes com esta ferramenta foram utilizados, maioritariamente, formatos de vídeo YUV encapsulados em ficheiros AVI. Foram usados dois *plugins*, um para construção das componentes gráficas (*Visual Swing for Eclipse*), e o *JFreeChart* para construir os gráficos com os resultados.

Esta aplicação foi testada com sucesso nas seguintes condições:

Windows XP; Windows Vista; Windows Seven;

Nas condições em que não foi testado, não tem nenhuma razão evidente, para que não funcione corretamente.

1.1 Requisitos

Para correto funcionamento da aplicação deve ser instalado o JMF e um JRE de 32bits. Não existe uma versão de 64bit do JMF por essa razão é necessário instalar um JRE de 32 bits nas máquinas que têm um sistema operativo de 64bit. Podendo coexistir ambos JRE (64bit e 32bit) no mesmo computador. Após instalação destes dois aplicativos pode-se correr o ficheiro executável (JAR) da aplicação.

2 Estrutura da interface gráfica

Na janela principal da interface gráfica está disponível uma barra de menus com três menus, como se pode observar na Figura 1. No menu ***Subjective Test*** pode-se obter uma breve descrição dos quatro métodos de avaliação subjetiva implementados (***Methods Info***). A aplicação disponibiliza um método *Single Stimulus* (ACR-HR), um método *Double Stimulus* (DSIS), um método de comparação (SCACJ) e um método de avaliação continua (AVCA), ver secção 6. Os três primeiros métodos estão nas recomendações ITU-T P.910 e ITU-R BT.500-12.

No submenu ***New Test*** pode-se efetuar a configuração de um teste de avaliação subjetiva, escolhendo inicialmente se queremos um dos métodos recomendados pela ITU ou o método de avaliação continua (AVCA). No submenu ***Load Test***, escolhe-se o

ficheiro de configuração gerado da opção anterior (**New Test**). Após validação do ficheiro de configuração será apresentado ao observador um painel para este introduzir os seus dados pessoais. De seguida iniciará o teste de avaliação subjetiva.

No menu **Data Analysis** são efetuados os cálculos e apresentação de resultados, com base nos ficheiros de pontuações escolhidos. Inicialmente seleciona-se se pretende visualizar os resultados de *Mean Opinion Score* (**Compute MOS**) ou os resultados do método de avaliação continua (**AVCA Results**). Na opção **Compute MOS** pode-se escolher o modo como se pretende que a rejeição de participantes com pontuações incoerentes seja calculada. Através do método de rejeição da ITU (**ITU-T method**) ou através do método da correlação (**Correlation Analysis**), ver secção 6.

No menu **Help** serão disponibilizadas informações sobre a ferramenta (**About**). Ao clicar no submenu **User Guide** será automaticamente visualizado um ficheiro em formato pdf a explicar o funcionamento da ferramenta (este documento).

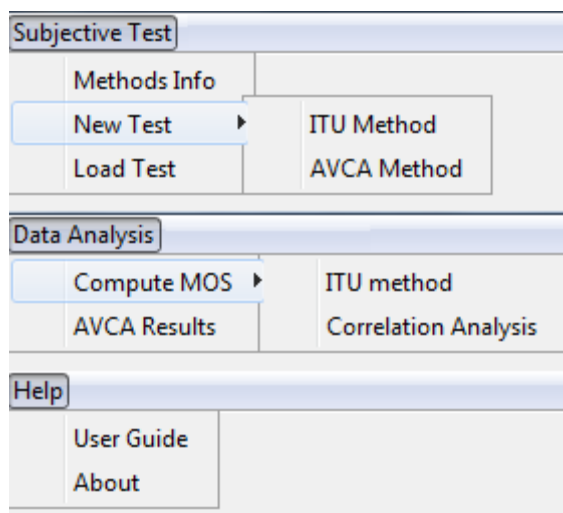


Figura 1: Menus da interface gráfica

Em seguida serão explicadas, detalhadamente, as seguintes partes da ferramenta: configuração de um novo teste, realização de um teste, apresentação de resultados.

3 Configuração de um novo teste

Um novo teste de avaliação subjetiva configura-se escolhendo a opção **New Test** do menu **Subjective Test**. Após escolha do tipo de teste que se pretende configurar, se um teste de avaliação continua (AVCA) ou um teste apresentado nas recomendações ITU, fica disponível um painel com as opções de configuração. Para uma configuração com um método recomendado pela ITU, ilustrada na Figura 2, este painel contém uma **combobox** que permite escolher o tipo de método de avaliação subjetiva, duas listas, **Video Group List** e **Display Order List**, onde serão adicionadas as sequências de vídeo escolhidas pelo organizador a figurar no teste.

Na primeira lista, as sequências de vídeo ficam agrupadas segundo ‘famílias’ ou grupos de vídeos. Estes conjuntos de vídeos são caracterizados pelo indentificador do grupo (G1, G2, ..., Gn). Estes vídeos são selecionados através do gestor de ficheiros apresentado após clicar no botão **Get Files**.

Os grupos de vídeos são adicionados à segunda lista através do botão **Add Group**. Quando se utilizam os métodos ACR-HR e DSIS, devido à sua metodologia, é necessário selecionar uma sequência de vídeo para ser a referência antes de se adicionar o grupo de vídeos à segunda lista. Para os métodos DSIS e SCACJ é necessário que cada grupo contenha pelo menos duas sequências de vídeo, pois os vídeos são exibidos aos pares.

Na segunda lista é apresentada a configuração final, onde as sequências estão organizadas pela ordem com que serão exibidas durante a realização do teste. Por exemplo, se o método ACR-HR for selecionado na *Display Order List*, cada sequência de vídeo deverá ser descrita pelo diretório da sequência de vídeo, seguido da identificação do grupo e identificador de referência ou condição de teste, “D:\Documents\AVIs\city.avi G1 R”, em que 1 é o número do grupo e ‘R’ significa que a sequência de vídeo é uma referência e ‘T’ significa que é de teste.

Se o método selecionado for o DSIS terá a seguinte configuração: identificação do grupo seguido do caminho para o ficheiro do vídeo de teste e caminho para o ficheiro do vídeo de referência. Se o método selecionado for o SCACJ será apresentado a identificação do grupo seguido do diretório da sequência de vídeo A e diretório da sequência de vídeo B. Na linha seguinte aparecerá a identificação do grupo seguido do caminho para o ficheiro do vídeo B e diretório da sequência de vídeo A, derivado a este método usar todas as combinações possíveis das sequências.

Como o tipo de método influencia a configuração dos ficheiros de teste, nomeadamente o modo como são criados e lidos. Após adicionar as sequências de vídeo contidas na *Video Group List* para a *Display Order List* deixa de ser possível mudar o tipo de método utilizado no teste (através da *combobox*), a não ser que a lista *Display Order List* fique novamente vazia.

As sequências de vídeo aparecem inicialmente agrupadas por grupos na *Display Order List*, mas a ordem com que serão visualizadas durante os testes pode ser gerada de uma forma pseudoaleatória ao selecionar o botão **Random**. Ao clicar no botão **Random** é escolhido aleatoriamente uma sequência de vídeo da lista e colocado novamente na *Display Order List*. O mesmo acontece para as restantes sequências de vídeo na lista com a seguinte regra - a sequência seguinte escolhida terá de ser de uma família de vídeos diferente da anterior e assim sucessivamente. Este espaçamento entre sequências do mesmo grupo/família é maior ou igual a metade do número de famílias total. Pode, no entanto, aparecer sequências de vídeo da mesma família seguidas, no fim da lista, caso os grupos de vídeo não tenham o mesmo número de sequências.

O gerador de números pseudoaleatórios pode ser iniciado com base num temporizador ou com base numa ‘semente’. Se for preenchido o campo **Seed id** com um número inteiro a ordem gerada para os elementos contidos na *Display Order List* será sempre a mesma desde que essa semente (um número inteiro) seja a mesma. Isto permite gerar a mesma ordem, para o mesmo conjunto de sequências, em momentos de tempo diferentes.

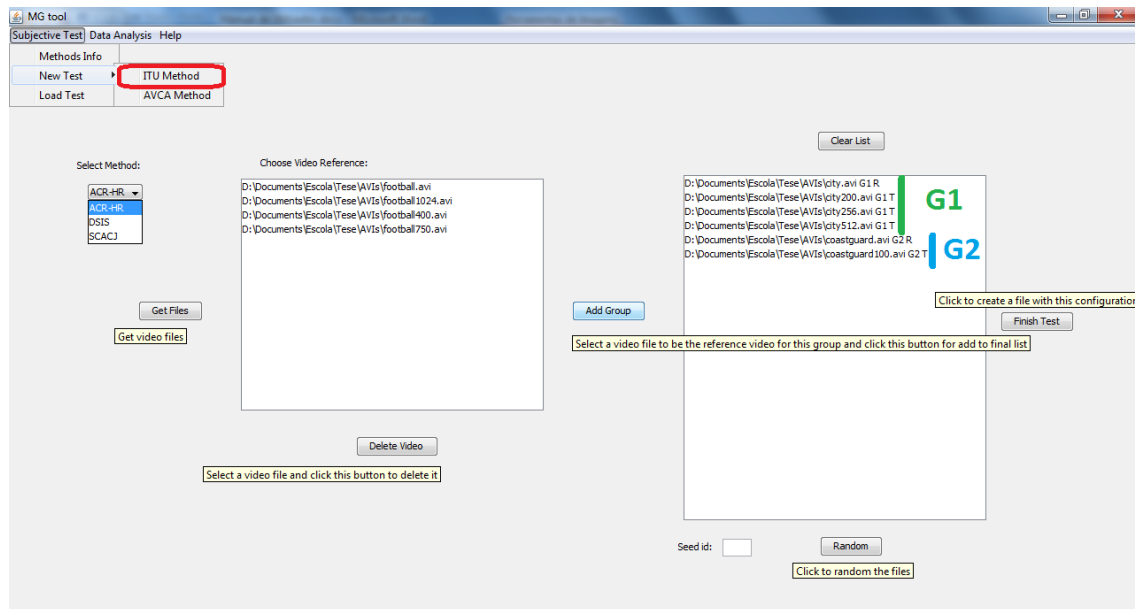


Figura 2: Configuração de um novo teste ACR-HR

Os vídeos contidos na primeira lista podem ser removidos da lista ao clicar no botão **Delete Video**, depois de se seleccionar a sequência de vídeo pretendida. Se o botão **Clear List** for premido, todas as sequências de vídeo contidas na segunda lista serão apagadas.

Ao pressionar os botões **Delete Video** ou **Add Group**, sem que exista qualquer sequência de vídeo seleccionada, ou se for clicado o botão **Add Group** com a *Video Group List* vazia, aparece uma mensagem de aviso a dizer que é necessário escolher pelo menos um vídeo. Se tentar clicar nos botões **Random**, **Clear List** ou **Finish Test** com a segunda lista vazia é mostrada uma mensagem de aviso a dizer que lista está vazia.

Antes de terminar a configuração do teste, e caso se pretenda fazer ajustes na ordem das sequências na *Display Order List*, pode-se utilizar o mecanismo de seleccionar uma linha desta lista e arrasta-la para a posição desejada (*drag and drop*). Este mecanismo facilita bastante nos casos em que se pretende alterar apenas algumas posições entre as sequências de vídeo sem alterar toda a ordem como faz o botão **Random**.

Após terminar a configuração do teste clica-se no botão **Finish Test** para criar o ficheiro de configuração do teste. Escolhe-se o directório onde queremos guardar e o nome do teste. É gerado um ficheiro de texto utilizando como título o nome inserido, com o seguinte conteúdo: nome do método, sequências de vídeo com os respectivos grupos e identificadores de sequências, conforme ilustrado na Figura 3, para o caso de um teste

com o método de avaliação ACR-HR. No final, a palavra ‘end’ assinala o fim do ficheiro de configuração. Esta palavra é posteriormente usada durante a realização do teste para se ficar a saber que este chegou ao fim.

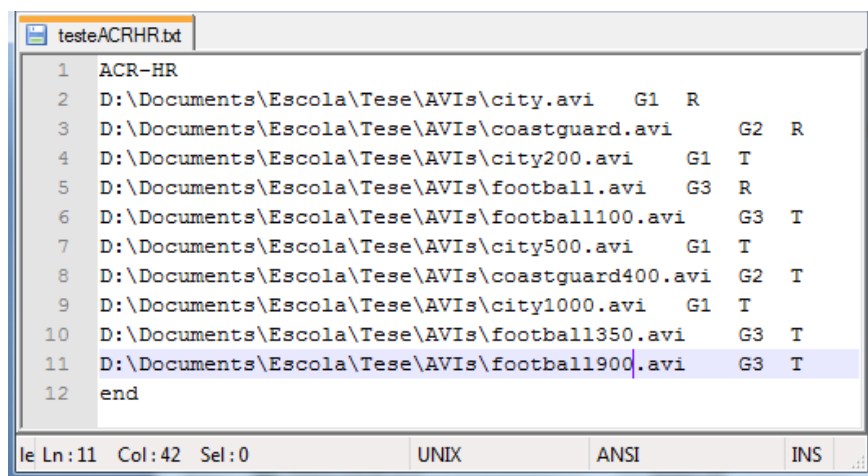


Figura 3: Ficheiro de configuração de um novo teste ACR-HR

Na Figura 4 é ilustrado o conjunto de parâmetros a configurar para um teste de avaliação continua (AVCA). Começa-se por seleccionar a sequência de vídeo, a figurar no teste, através do botão **Get File**. Os tempos de artefactos (perturbações no vídeo) são adicionados à lista final de tempos através do botão **Add time**. Estes tempos devem estar em segundos. Para eliminar um tempo da lista de tempos basta selecciona-lo e clicar no botão **Delete time**.

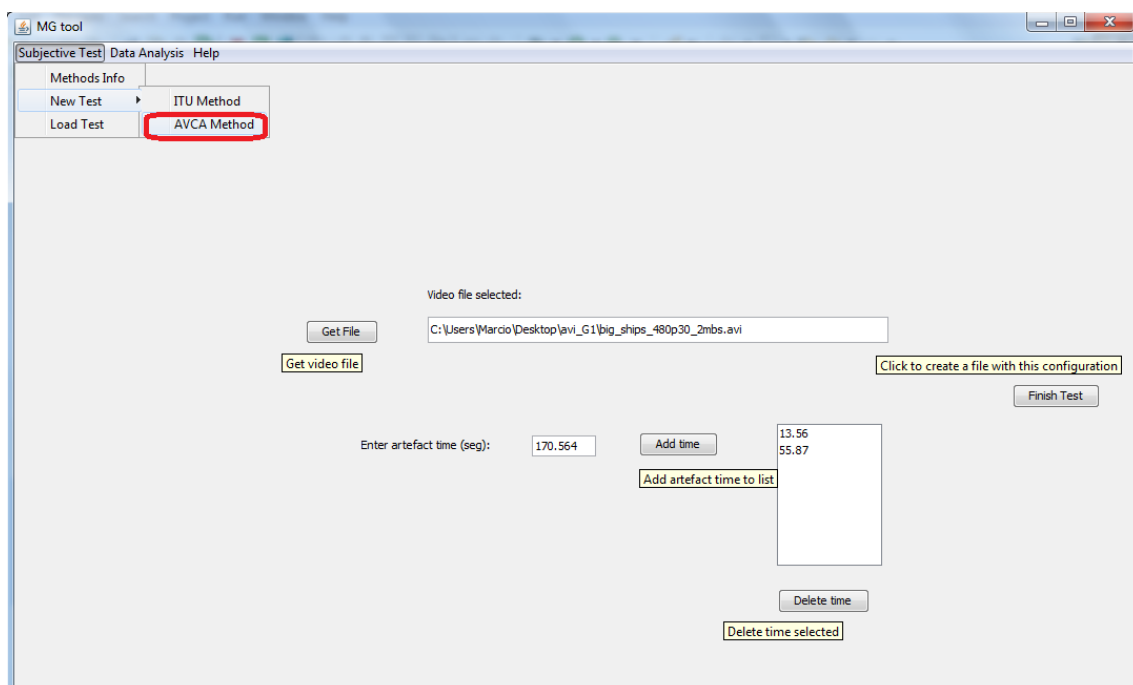


Figura 4: Configuração de um novo teste AVCA

Após terminar a configuração do teste clicka-se no botão **Finish Test** para criar o ficheiro de configuração do teste. Escolhe-se o diretório onde queremos guardar o ficheiro e o

nome do teste. É gerado um ficheiro de texto utilizando como título o nome inserido, com o seguinte conteúdo: nome do método, sequência de vídeo, número de tempos de artefactos introduzidos e respetivos tempos, conforme ilustrado na Figura 5. No final, a palavra ‘end’ assinala o fim do ficheiro de configuração.

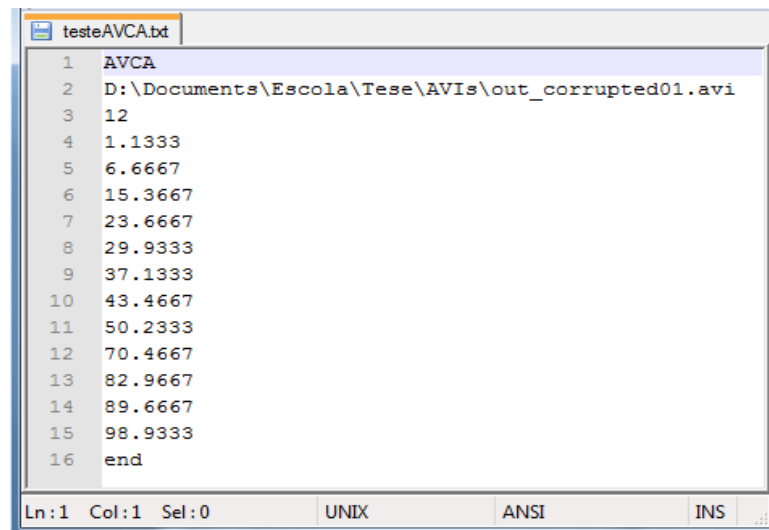


Figura 5: Ficheiro de configuração de um novo teste AVCA

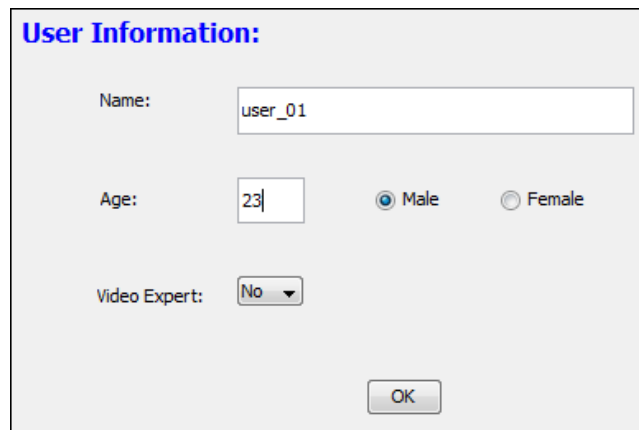
Quando se faz a configuração de um teste num PC e depois se pretende realizá-lo noutro PC, tem que se ter em atenção os diretórios dos vídeos. Se houver diretórios diferentes de um PC para outro, na altura em que se faz o **Load Test** será mostrada uma mensagem de erro. Para contornar este problema pode-se utilizar um dispositivo externo ou exatamente o mesmo diretório para as sequências de vídeo.

4 Realização de um teste

Conforme dito anteriormente, para a realização de um teste de avaliação subjetiva escolhe-se a opção **Load Test** no menu **Subjective Test**. É mostrado um gestor de ficheiros para escolher o ficheiro de texto com o teste pretendido.

Para garantir que todas as sequências de vídeo podem ser lidas corretamente é feita uma verificação aos caminhos dos diretórios onde estão armazenados as sequências de vídeo do ficheiro escolhido. Se algum destes caminhos estiver incorreto e que dê origem a erros na visualização do vídeo, é mostrada uma mensagem onde informa que o respetivo ficheiro de vídeo não pode ser lido. Se não existir nenhum erro no carregamento do ficheiro é disponibilizado um painel para o participante inserir os seus dados.

Neste painel, ilustrado na Figura 6, o utilizador deverá preencher os campos **nome**, **idade**, **género** e **tipo de participante** (se é especialista em vídeo ou não). Depois de feitas as verificações a estes campos – **nome** não pode estar sem conteúdo e **idade** tem de ser um número inteiro entre 1 e 99 – é dado início à sessão de avaliação subjetiva da qualidade do vídeo. O nome do participante vai ser usado para o nome do ficheiro gerado com os resultados provenientes da sessão.

A dialog box titled "User Information:" in blue text. It contains three input fields: "Name:" with the text "user_01", "Age:" with the number "23", and "Video Expert:" with a dropdown menu showing "No". To the right of the "Age:" field are two radio buttons labeled "Male" (selected) and "Female". At the bottom center is an "OK" button.

User Information:

Name: user_01

Age: 23 ☒ Male ☐ Female

Video Expert: No

OK

Figura 6: Painel de informação sobre o participante

Para evitar que os participantes se distraiam e percam a concentração na avaliação das sequências de vídeo, os menus da barra de ferramentas deixam de estar disponíveis após serem guardadas as informações relativas ao participante. Caso essas opções estivessem disponíveis e fossem selecionadas iriam cancelar a avaliação feita até ao momento. Os menus voltam a ficar disponíveis quando o teste terminar.

Assim é dado início à sessão de avaliação subjetiva que começa com uma breve explicação sobre o tipo de método utilizado no teste (exemplo na Figura 7 para um teste com o método DSIS), e de que forma as sequências de vídeo vão ser exibidas. Durante esta breve introdução aborda-se também a escala de pontuações que vai ser usada para avaliação da qualidade dos vídeos. Todos estes parâmetros dependem do tipo de método utilizado no teste.

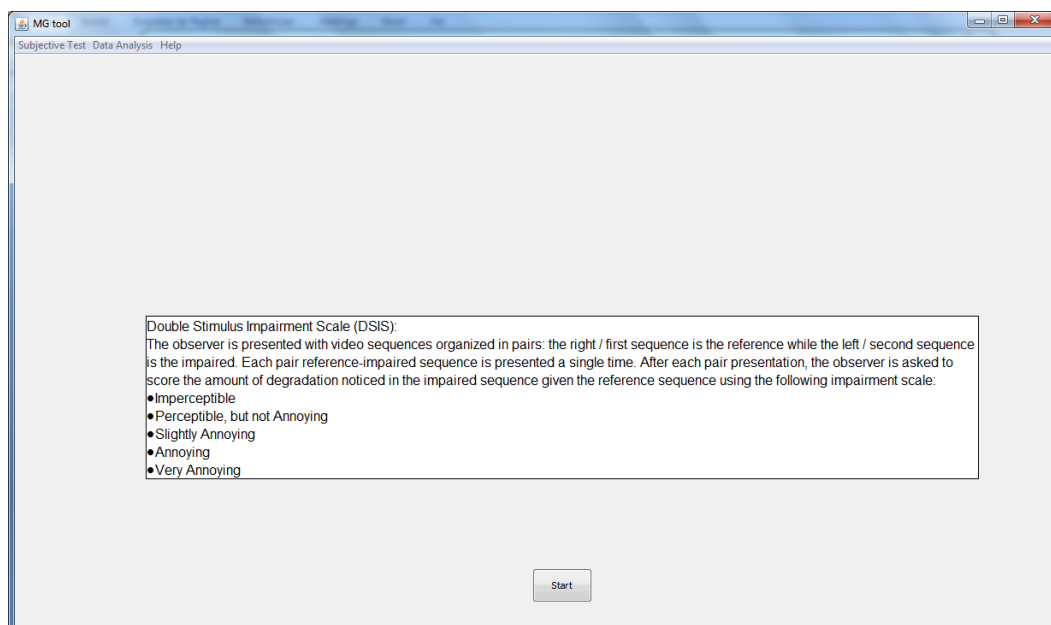


Figura 7: Descrição do método DSIS

O participante clica no botão **Start** para dar início à visualização das sequências de vídeo. Enquanto o vídeo está a ser visualizado, Figura 8, o botão **Start** desaparece e só

volta a estar visível depois de efetuada a avaliação de cada sequência de vídeo exibida no teste.

Dois segundos após a sequência de vídeo ter terminado fica visível a escala de avaliação, ilustrado na Figura 9. Na janela correspondente à escala de avaliação apenas é permitido escolher uma das pontuações disponíveis. Depois de visualizado o vídeo e pontuada a sua qualidade, o participante clica no botão **Start**, para iniciar a próxima sequência de vídeo, como descrito acima, é também neste momento que é guardado num ficheiro a pontuação da sequência de vídeo visualizada e respetivo nome. Este procedimento repete-se sucessivamente até aparecer uma janela a dizer que o teste terminou.

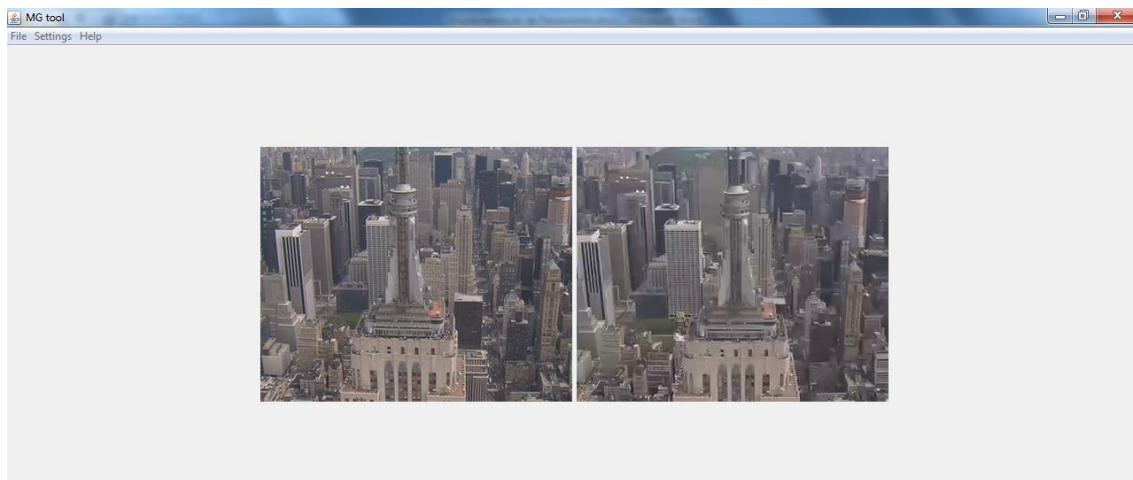


Figura 8: Teste de avaliação do método DSIS

De acordo com a recomendação ITU-R BT.500-12, quando se utilizam os métodos DSIS ou SCACJ, as sequências de vídeo devem ser mostradas em pares lado a lado, quando as dimensões do ecrã o permitam. No caso particular do método DSIS, a sequência de referência é o primeiro vídeo a ser visualizado ou, no caso de as sequências serem exibidas em simultâneo, é o vídeo apresentado no lado esquerdo do ecrã.

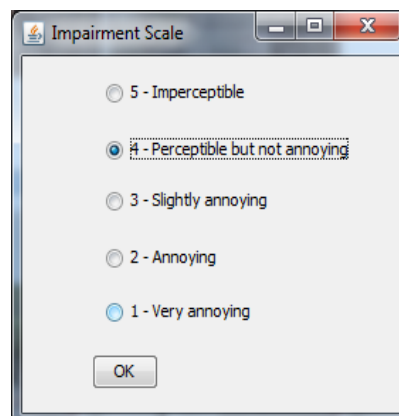
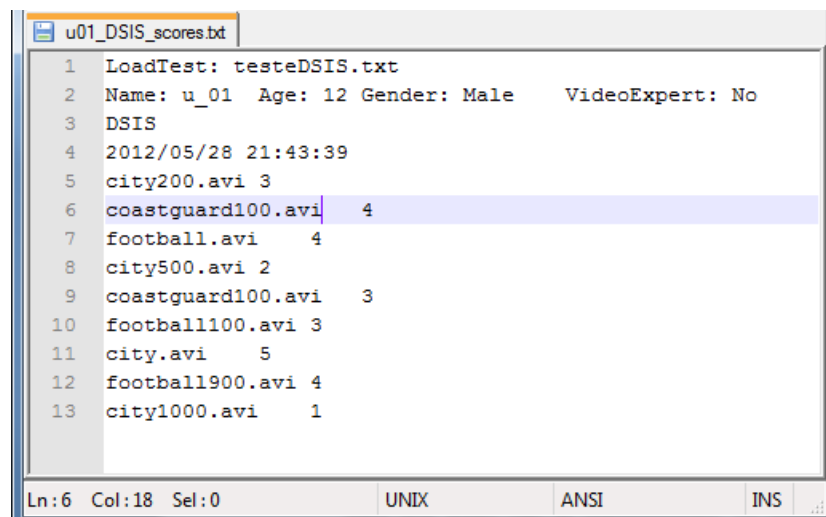


Figura 9: Escala de avaliação do método DSIS

Terminado o teste de avaliação subjetiva, as pontuações dadas pelo participante ficam guardadas num ficheiro de texto. O ficheiro é gravado com o seguinte nome “**nomeParticipante_nomeMetodoAvaliaçãoSubjetiva_scores.txt**” e contém o nome do ficheiro de configuração do teste, a informação relativa ao participante (nome, idade, género e tipo de participante), o nome do método de avaliação subjetiva, a data e hora em que se efetuou o teste, e finalmente o nome das sequências de vídeo avaliadas e respetiva pontuação, como ilustrado na Figura 10.

No caso do método de avaliação subjetiva escolhido ser o SCACJ, a única diferença no formato do ficheiro das pontuações é ter os nomes do par de sequências e não apenas a sequência de vídeo avaliada. Esta opção justifica-se pelo facto de, neste método, o par de sequências de vídeo avaliadas resultar de combinações entre os ficheiros de vídeo pertencentes a cada grupo.

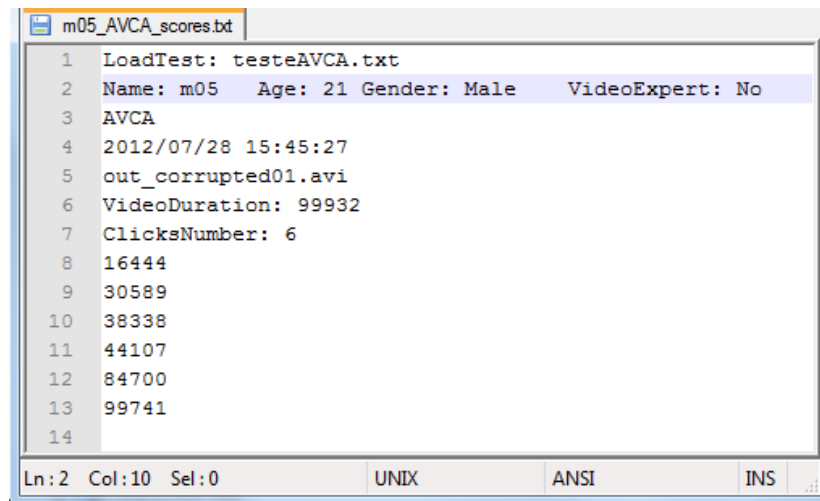
Para o método de avaliação continua AVCA após a informação sobre a data e hora da realização do teste é descrito o nome do vídeo, a sua duração, quantas vezes o participante clicou durante o teste e os respetivos tempos desses *clicks* (em milissegundos), conforme ilustrado na Figura 11.



```
u01_DSIS_scores.txt
1 LoadTest: testeDSIS.txt
2 Name: u_01 Age: 12 Gender: Male VideoExpert: No
3 DSIS
4 2012/05/28 21:43:39
5 city200.avi 3
6 coastguard100.avi 4
7 football.avi 4
8 city500.avi 2
9 coastguard100.avi 3
10 football100.avi 3
11 city.avi 5
12 football900.avi 4
13 city1000.avi 1
```

Ln:6 Col:18 Sel:0 UNIX ANSI INS

Figura 10: Ficheiro de pontuações de um participante com o método DSIS



```
1 LoadTest: testeAVCA.txt
2 Name: m05 Age: 21 Gender: Male VideoExpert: No
3 AVCA
4 2012/07/28 15:45:27
5 out_corrupted01.avi
6 VideoDuration: 99932
7 ClicksNumber: 6
8 16444
9 30589
10 38338
11 44107
12 84700
13 99741
14
```

Figura 11: Ficheiro de pontuações de um participante com o método AVCA

Os ficheiros com as pontuações dos participantes devem ser gravados com nomes diferentes para que estes não se sobreponham. Os ficheiros de pontuações são guardados na mesma diretoria do ficheiro de configuração do teste.

5 Apresentação de resultados

Os resultados são obtidos através da análise dos ficheiros de pontuações escolhidos. Os ficheiros selecionados devem ter em comum o mesmo ficheiro de configuração do teste, para que seja possível efetuar os cálculos corretamente. Caso isso não aconteça é mostrada uma mensagem de erro.

No menu **Data Analysis** seleciona-se pelo tipo de resultados que se pretende. Resultados para testes com o método de avaliação contínua da qualidade do vídeo (**AVCA Results**) ou resultados para obter valores de MOS (**Compute MOS**). Neste último submenu escolhe-se o tipo de método para a rejeição de pontuações incoerentes por parte dos observadores. De seguida os ficheiros com as pontuações dos participantes são selecionados a partir de um gestor de ficheiros.

O cálculo de *Mean Opinion Score* (MOS) e rejeição de participantes com pontuações incoerentes, são feitos de duas formas distintas (**ITU method** e **Correlation Analysis**). A apresentação dos resultados, em forma de tabela, depende do tipo de método de avaliação subjetiva usado no teste. Caso se escolha a opção **ITU method** e na realidade o nome do método contido no ficheiro seja o AVCA, ou vice-versa, é mostrada uma janela de erro.

Para um método ITU as colunas da tabela são constituídas pelos seguintes campos:

- nome das sequências de vídeo;
- nome dos participantes, em que o total de participantes corresponde ao número de ficheiros selecionados anteriormente;
- *RawMOS*, baseado nas pontuações de todos os participantes;

- desvio padrão de *RawMOS*;
- MOS, baseado nas pontuações de todos os participantes válidos;
- desvio padrão de MOS;
- vídeo de referência – para cada sequência de vídeo indica se é um vídeo de referência ou não; (apenas para ACR-HR)
- *Differential Mean Opinion Score* (DMOS); (apenas para ACR-HR)

Conforme ilustrado na Figura 12, os nomes das sequências de vídeo exibidas durante os testes de avaliação subjetiva aparecem na primeira coluna da esquerda. As pontuações de cada participante, para a sequência de vídeo correspondente, surgem nas colunas imediatamente seguintes. Para o cálculo do *RawMOS* e do seu desvio padrão estão abrangidos todos os participantes. Para o cálculo dos valores de MOS e respectivos desvio padrão apenas se considera os participantes válidos, ou seja, os participantes que não apresentem pontuações incoerentes. As pontuações atribuídas pelos participantes que foram rejeitados para o cálculo do MOS ficam com fundo amarelo. Os resultados são arredondados para 2 casas decimais.

Video Sequence	m01	m02	m03	m04	m05	m06	m07	m08	m09	RawMOS	SD RawMOS	MOS	SD MOS	Video Reference	DMOS
soccer_480p30_3mbps_corrupted_plr_3.2325_avi	2	2	1	2	2	1	1	2	2	1.67	0.5	1.67	0.5	NO	2.89
soccer_480p30_3mbps_corrupted_plr_1.7177_avi	1	1	1	2	1	1	1	1	1	1.11	0.6	1.11	0.6	NO	3.45
soccer_480p30_3mbps_corrupted_plr_1.3832_avi	2	2	2	3	2	2	2	2	2	2.11	0.69	2.11	0.69	NO	2.45
snow_mnt_480p30_3mbps_corrupted_plr_0.60796_avi	4	3	4	1	4	4	4	3	3	3.33	1.21	3.33	1.21	NO	1.67
old_town_cross_480p25_2mbps_corrupted_plr_0.563_avi	4	5	4	1	4	4	4	5	5	4.0	1.72	4.0	1.72	NO	0.56
station_480p25_2mbps_corrupted_plr_0.22712_avi	4	4	5	1	5	5	5	4	4	4.11	2.14	4.11	2.14	NO	-1.11
sunflower_480p25_2mbps_corrupted_plr_7.7564_avi	1	1	1	5	1	1	1	1	1	1.44	2.52	1.44	2.52	NO	3.12
park_joy_480p25_5mbps_avi	5	5	5	1	5	5	5	5	5	4.56	2.85	4.56	2.85	YES	0.0
raven_480p30_2mbps_corrupted_plr_5.1629_avi	3	2	3	3	3	3	3	2	2	2.67	2.9	2.67	2.9	NO	1.89
snow_mnt_480p30_3mbps_corrupted_plr_2.2527_avi	3	3	3	3	3	3	3	3	3	3.0	2.9	3.0	2.9	NO	2.0
big_ships_480p30_2mbps_avi	5	5	5	1	5	5	5	5	5	4.56	3.19	4.56	3.19	YES	0.0
sunflower_480p25_2mbps_corrupted_plr_2.4583_avi	2	2	2	2	2	2	2	2	2	2.0	3.19	2.0	3.19	NO	2.56
station_480p25_2mbps_corrupted_plr_1.9709_avi	3	3	3	3	3	3	3	3	3	3.0	3.19	3.0	3.19	NO	0.0
raven_480p30_2mbps_corrupted_plr_1.7177_avi	3	4	3	3	3	3	3	4	4	3.33	3.23	3.33	3.23	NO	1.23
old_town_cross_480p25_3mbps_corrupted_plr_2.5564_avi	2	2	2	2	2	2	2	2	2	2.0	3.23	2.0	3.23	NO	2.56
park_joy_480p25_5mbps_corrupted_plr_3.8627_avi	2	2	2	2	2	2	2	2	2	2.0	3.23	2.0	3.23	NO	2.56
soccer_480p30_3mbps_corrupted_plr_1.3028_avi	5	5	5	1	5	5	5	5	5	4.56	3.49	4.56	3.49	NO	0.0
sunflower_480p25_2mbps_corrupted_plr_0.53605_avi	4	4	4	1	4	4	4	4	4	3.67	3.63	3.67	3.63	NO	0.89
station_480p25_2mbps_corrupted_plr_0.21373_avi	5	5	5	1	5	5	5	5	5	4.56	3.87	4.56	3.87	NO	-1.56
snow_mnt_480p30_3mbps_corrupted_plr_0.59624_avi	5	5	5	1	5	5	5	5	5	4.56	4.09	4.56	4.09	NO	0.44
big_ships_480p30_2mbps_corrupted_plr_5.1629_avi	2	2	2	2	2	2	2	2	2	2.0	4.09	2.0	4.09	NO	2.56
park_joy_480p25_5mbps_corrupted_plr_0.21373_avi	2	2	2	2	2	2	2	2	2	2.0	4.09	2.0	4.09	NO	2.56
raven_480p30_2mbps_avi	5	5	5	1	5	5	5	5	5	4.56	4.3	4.56	4.3	YES	0.0
old_town_cross_480p25_2mbps_corrupted_plr_7.0338_avi	2	2	2	2	2	2	2	2	2	2.0	4.3	2.0	4.3	NO	2.56
station_480p25_2mbps_corrupted_plr_1.9848_avi	3	3	3	3	3	3	3	3	3	3.0	4.3	3.0	4.3	NO	0.0
snow_mnt_480p30_3mbps_avi	5	5	5	1	5	5	5	5	5	4.56	4.51	4.56	4.51	YES	0.44
sunflower_480p25_2mbps_corrupted_plr_7.0338_avi	1	1	1	5	1	1	1	1	1	1.44	4.7	1.44	4.7	NO	3.12
big_ships_480p30_2mbps_corrupted_plr_0.19048_avi	4	4	4	1	4	4	4	4	4	3.67	4.8	3.67	4.8	NO	0.89
raven_480p30_2mbps_corrupted_plr_5.0147_avi	3	3	3	3	3	3	3	3	3	3.0	4.8	3.0	4.8	NO	1.56
park_joy_480p25_5mbps_corrupted_plr_1.5705_avi	2	2	2	2	2	2	2	2	2	2.0	4.8	2.0	4.8	NO	2.56
station_480p25_2mbps_avi	5	5	5	1	5	5	5	5	5	4.56	4.99	4.56	4.99	YES	-1.56
old_town_cross_480p25_3mbps_corrupted_plr_7.7564_avi	3	3	3	3	3	3	3	3	3	3.0	4.99	3.0	4.99	NO	1.56
big_ships_480p30_2mbps_corrupted_plr_0.16803_avi	4	4	4	4	4	4	4	4	4	4.0	4.99	4.0	4.99	NO	0.56
snow_mnt_480p30_3mbps_corrupted_plr_7.8129_avi	2	2	2	2	2	2	2	2	2	2.0	4.99	2.0	4.99	NO	3.0
soccer_480p30_3mbps_avi	5	5	5	5	5	5	5	5	5	5.0	4.99	5.0	4.99	YES	-0.44
park_joy_480p25_5mbps_corrupted_plr_7.0338_avi	1	1	1	1	1	1	1	1	1	1.0	4.99	1.0	4.99	NO	3.56
sunflower_480p25_2mbps_corrupted_plr_0.563_avi	4	4	4	4	4	4	4	4	4	4.0	4.99	4.0	4.99	NO	0.56
raven_480p30_2mbps_corrupted_plr_1.7061_avi	3	3	3	3	3	3	3	3	3	3.0	4.99	3.0	4.99	NO	1.56
big_ships_480p30_2mbps_corrupted_plr_5.0147_avi	1	1	1	1	1	1	1	1	1	1.0	4.99	1.0	4.99	NO	3.56
old_town_cross_480p25_2mbps_corrupted_plr_2.4583_avi	1	1	1	1	1	1	1	1	1	1.0	4.99	1.0	4.99	NO	3.56

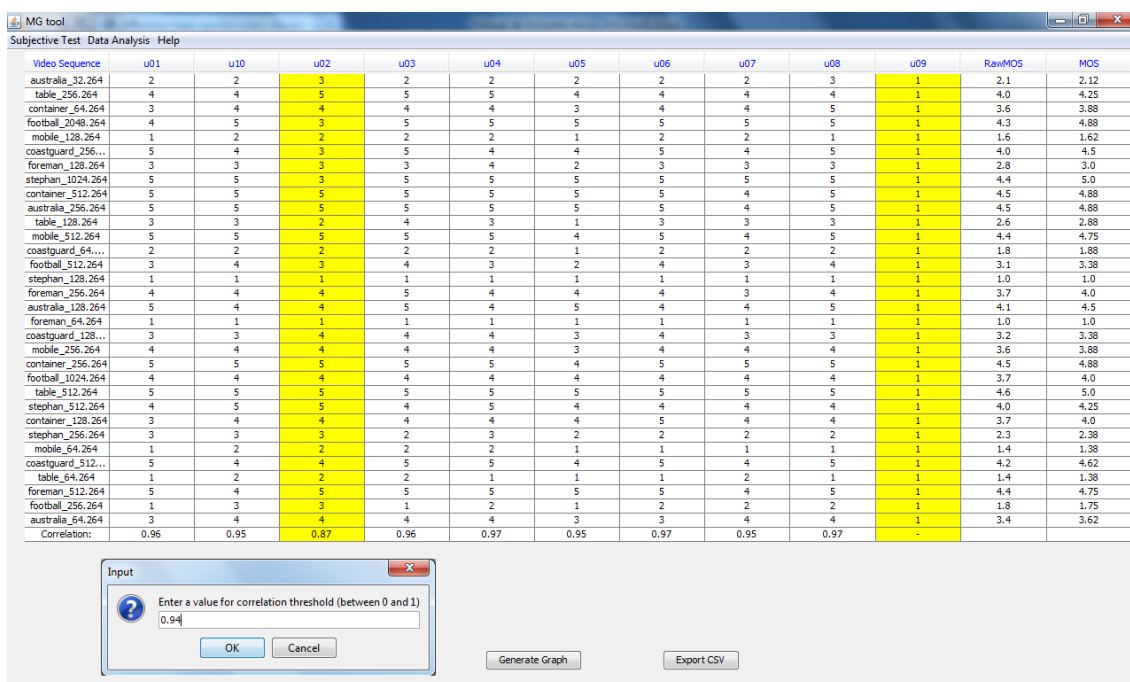
Figura 12: Tabela de resultados utilizando a metodologia ACR-HR com método de rejeição ITU

Após visualização da tabela com os resultados dos observadores ficam disponíveis dois novos botões. Um para gerar um gráfico (**Generate Graph**) com os resultados e outro para exportar os resultados para um ficheiro CSV (**Export CSV**).

Como cálculo de rejeição de resultados para o método da correlação é diferente, as colunas na tabela também são diferentes dos métodos da ITU. A primeira coluna contém o nome das sequências de vídeo, as colunas seguintes têm as pontuações de cada observador. Nas duas últimas colunas são visíveis os valores de *RawMOS* e MOS respetivamente. A última linha da tabela é disponibilizado o valor do coeficiente da

correlação para cada observador. O utilizador é que define o valor de referência para coeficiente de correlação a partir do qual os participantes devem ser rejeitados.

Conforme ilustrado na Figura 13 os observadores com valores de coeficiente de correlação inferiores ao definido são rejeitados, colunas a amarelo. Para este exemplo definiu-se um coeficiente de correlação de 0.94. Caso um observador atribua a mesma pontuação a todas as sequências de vídeo definiu-se que as suas pontuações também deveriam ser rejeitadas. E o valor apresentado para o seu coeficiente de correlação será ‘-’ (indefinido).



The screenshot shows the 'MG tool' window with a menu bar (Subjective Test, Data Analysis, Help) and a table of results. The table has columns for Video Sequence, u01, u10, u02, u03, u04, u05, u06, u07, u08, u09, RawMOS, and MOS. The 'u' columns contain numerical ratings, while 'RawMOS' and 'MOS' contain calculated values. Some cells in the 'u' columns are highlighted in yellow, indicating a correlation coefficient below the threshold. An 'Input' dialog box is open, prompting the user to 'Enter a value for correlation threshold (between 0 and 1)' with '0.94' entered. Below the dialog are buttons for 'OK', 'Cancel', 'Generate Graph', and 'Export CSV'.

Video Sequence	u01	u10	u02	u03	u04	u05	u06	u07	u08	u09	RawMOS	MOS
australia_32.264	2	2	3	2	2	2	2	2	3	1	2.1	2.12
table_256.264	4	4	5	5	5	4	4	4	4	1	4.0	4.25
container_64.264	3	4	4	4	4	3	4	4	5	1	3.6	3.88
football_2048.264	4	5	3	5	5	5	5	5	5	1	4.3	4.88
mobile_128.264	1	2	2	2	2	1	2	2	1	1	1.6	1.62
coastguard_256...	5	4	3	5	4	4	5	4	5	1	4.0	4.5
foreman_128.264	3	3	3	3	4	2	3	3	3	1	2.8	3.0
stephan_1024.264	5	5	3	5	5	5	5	5	5	1	4.4	5.0
container_512.264	5	5	5	5	5	5	5	4	5	1	4.5	4.88
australia_256.264	5	5	5	5	5	5	5	4	5	1	4.5	4.88
table_128.264	3	3	2	4	3	1	3	3	3	1	2.6	2.88
mobile_512.264	5	5	5	5	5	4	5	4	5	1	4.4	4.75
coastguard_64...	2	2	2	2	2	1	2	2	2	1	1.8	1.88
football_512.264	3	4	3	4	3	2	4	3	4	1	3.1	3.38
stephan_128.264	1	1	1	1	1	1	1	1	1	1	1.0	1.0
foreman_256.264	4	4	4	5	4	4	4	3	4	1	3.7	4.0
australia_128.264	5	4	4	5	4	5	4	4	5	1	4.1	4.5
foreman_64.264	1	1	1	1	1	1	1	1	1	1	1.0	1.0
coastguard_128...	3	3	4	4	4	3	4	3	3	1	3.2	3.38
mobile_256.264	4	4	4	4	4	3	4	4	4	1	3.6	3.88
container_256.264	5	5	5	5	5	4	5	5	5	1	4.5	4.88
football_1024.264	4	4	4	4	4	4	4	4	4	1	3.7	4.0
table_512.264	5	5	5	5	5	5	5	5	5	1	4.6	5.0
stephan_512.264	4	5	5	4	5	4	4	4	4	1	4.0	4.25
container_128.264	3	4	4	4	4	4	5	4	4	1	3.7	4.0
stephan_256.264	3	3	3	2	3	2	2	2	2	1	2.3	2.38
mobile_64.264	1	2	2	2	2	1	1	1	1	1	1.4	1.38
coastguard_512...	5	4	4	5	5	4	5	4	5	1	4.2	4.62
table_64.264	1	2	2	2	1	1	1	2	1	1	1.4	1.38
foreman_512.264	5	4	5	5	5	5	5	4	5	1	4.4	4.75
football_256.264	1	3	3	1	2	1	2	2	2	1	1.8	1.75
australia_64.264	3	4	4	4	4	3	3	4	4	1	3.4	3.62
Correlation:	0.96	0.95	0.87	0.96	0.97	0.95	0.97	0.95	0.97	-		

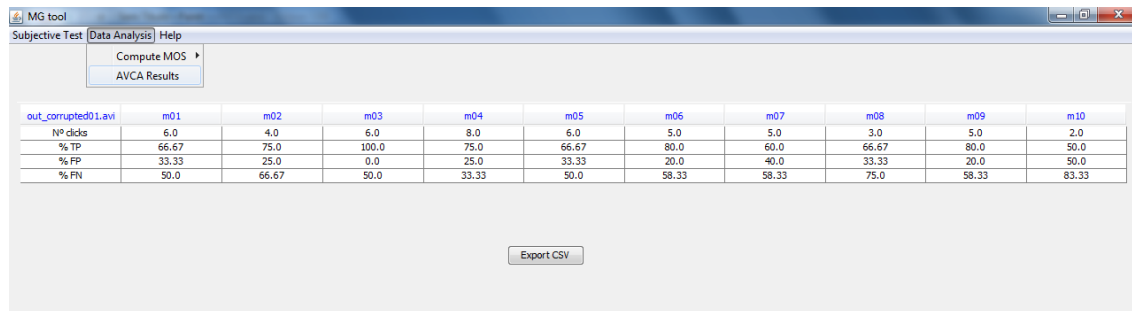
Figura 13: Tabela de resultados utilizando metodologia DSIS com método de rejeição Correlação

A apresentação de resultados para o método de avaliação contínua (**AVCA Results**) difere bastante da anterior. Para este método interessa saber o número de vezes que o observador assinalou a área de visualização do vídeo em relação ao número de artefactos introduzidos no vídeo. E analisar a percentagem de vezes que este clicou acertadamente (TP), clicou erradamente (FP) e percentagem de vezes que este não clicou e deveria ter clicado (FN), ou seja, não viu os artefactos mas eles estavam lá.

Para o cálculo destas percentagens é contabilizado o tempo de reação do observador – tempo que o observador demora a clicar no vídeo após ter visualizado o artefacto. Após análise de alguns testes efetuados para testar o tempo de reação das pessoas definiu-se que este tempo seria igual a 1,5 segundo. Um exemplo destes resultados é apresentado na Figura 14.

Durante a visualização do vídeo para este método é importante que não exista problemas de processamento no computador. Porque se houver paragens na sequência de vídeo devido a processamento no computador o observador pode assumir que era um

artefacto devido a perdas de codificação e clicar no botão do rato. No cálculo de resultados este *click* entra na percentagem dos falsos positivos (FP).



out_corrupted01.avi	m01	m02	m03	m04	m05	m06	m07	m08	m09	m10
Nº clicks	6.0	4.0	6.0	8.0	6.0	5.0	5.0	3.0	5.0	2.0
% TP	66.67	75.0	100.0	75.0	66.67	80.0	60.0	66.67	80.0	50.0
% FP	33.33	25.0	0.0	25.0	33.33	20.0	40.0	33.33	20.0	50.0
% FN	50.0	66.67	50.0	33.33	50.0	58.33	58.33	75.0	58.33	83.33

Figura 14: Tabela de resultados com o método AVCA

6 Avaliação subjetiva

Absolute Category Rating with Hidden Reference (ACR-HR) – Neste método, as sequências de vídeo são apresentadas ao participante uma de cada vez e este faz a avaliação entre exibições consecutivas, com base numa escala de qualidade representada na Tabela 1. As pontuações de qualidade são geralmente expressas por valores de *Differential Mean Opinion Scores* (DMOS), calculados com base nas diferenças entre as pontuações dadas ao vídeo degradado e as dadas ao vídeo de referência. O vídeo de referência é mostrado entre as restantes sequências de vídeos, sem que seja informado ao observador que está a pontuar o vídeo de referência. Essa é a razão pela qual o método se chama “Referência Escondida”.

Double Stimulus Impairment Scale (DSIS) – Neste método é apresentado ao participante uma referência (vídeo original) e o mesmo vídeo mas sob uma determinada condição de teste (vídeo degradado). O participante atribui uma pontuação ao vídeo degradado, que é feita tendo em conta o conteúdo do vídeo de referência. As pontuações atribuídas às sequências de vídeo devem ser conforme a Tabela 2. Se o avaliador não notar nenhum artefacto no vídeo de teste deve dar nota 5. Quer isto dizer que a diferença entre o vídeo original (referência) e o vídeo sob condição de teste é impercetível. Esta referência deve ter uma qualidade superior aos vídeos sob condições de teste. Este método tem duas variantes, nesta aplicação foi implementado a primeira variante – cada par referência/condição de teste é mostrado uma única vez, pedindo-se ao avaliador para votar após a exibição de cada par.

Stimulus Comparison Adjectival Categorical Judgement (SCACJ) – Neste método cada par corresponde a duas condições de teste distintas. A avaliação é feita pedido aos observadores que indiquem qual das condições de teste de cada par é a que possui melhor qualidade. Geralmente este teste é feito utilizando todos os pares de condições de teste possíveis, e deve-se assegurar que as sequências de cada par são visualizadas em ambas as ordens possíveis. Com base nas classificações apresentadas na Tabela 3, o observador irá indicar se a segunda sequência de vídeo tem melhor, pior ou a mesma qualidade da primeira e quão melhor ou pior é essa qualidade.

Artifact Visibility Continuous Assessment (AVCA) – Neste método é apresentado ao observador uma sequência de vídeo longa (maior que 2 minutos). Esta sequência de vídeo consiste num conjunto de cenas de vários vídeos. Durante a apresentação da sequência de vídeo o observador sempre que visualizar um artefacto (perturbação na imagem) deve assinalá-lo com o botão do rato na área do vídeo. Não existe uma escala de avaliação de qualidade pois esta é feita de forma continua durante a apresentação do vídeo.

Tabela 1: Escala de Qualidade

5	Excelente
4	Boa
3	Razoável
2	Pobre
1	Má

Tabela 2: Escala de Artefactos

5	Impercetível
4	Percetível
3	Ligeiramente Incómodo
2	Incómodo
1	Muito Incómodo

Tabela 3: Escala de Comparação SCACJ

-3	Muito Pior
-2	Pior
-1	Ligeiramente Pior
0	Igual
1	Ligeiramente Melhor
2	Melhor
3	Muito Melhor

ITU-T method – As pontuações dos observadores são processadas de acordo com a recomendação ITU-R BT.500-12. As pontuações fora do intervalo de confiança podem ser rejeitadas.

Correlation method – A rejeição de resultados incoerentes é feita com base na correlação entre as pontuações do observador e a média de todos os observadores. Inicialmente é escolhido um valor para o coeficiente de correlação, entre 0 e 1. Todos observadores que obtenham um coeficiente de correlação inferior ao definido terão as suas pontuações rejeitadas.