

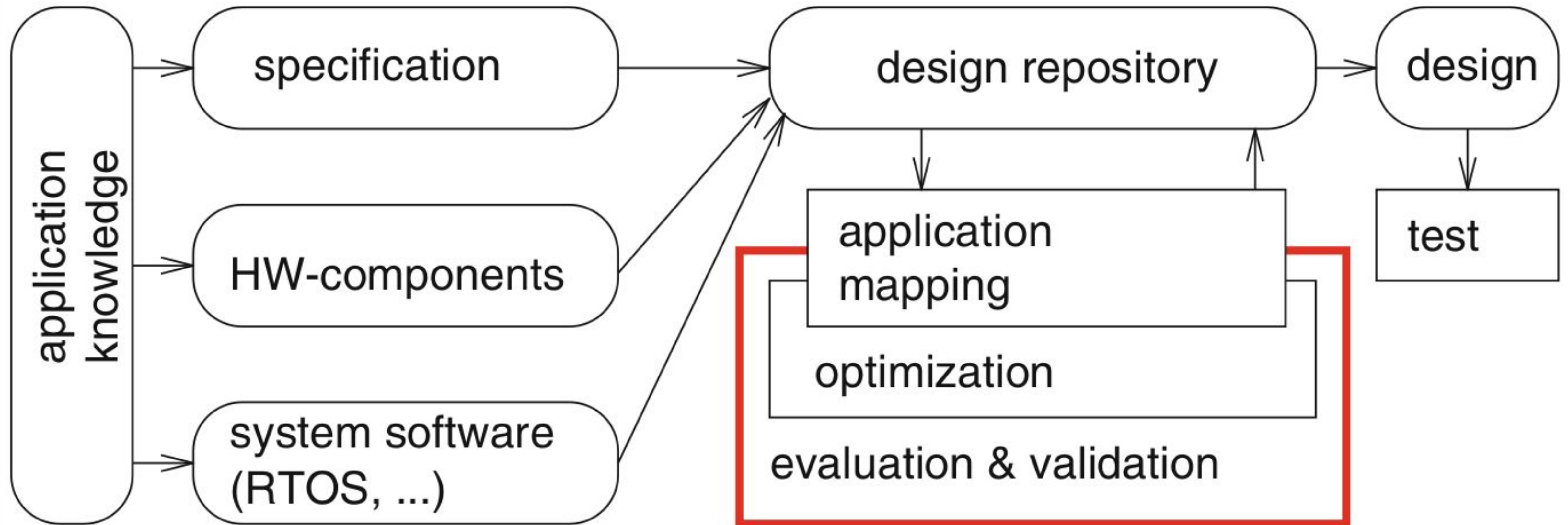
# Projeto de Sistemas Embarcados

IBM3119

Prof. Rigel Fernandes



# Avaliação



# Avaliação

- A etapa de avaliação antecede a etapa de otimização
- Quantificar o desempenho
- Devem ser realizadas durante os etapas de um projeto
- Essa aula ajudará a calcular medidas de desempenho dos projetos (ou parte do projeto)
- Definition: evaluation is the process of computing quantitative information of some key characteristics (or “objectives”) of a certain (possibly partial) design

# Avaliação

- Exemplos de medidas de interesse
  - medir o erro/acurácia
  - medir o erro absoluto
  - erro médio absoluto
  - erro médio com desvio padrão
  - erro quadrático médio com desvio padrão
  - Além disso, pode ser necessário verificar o erro para melhor entendimento
  - medir o tempo de execução (médio, desvio padrão, melhor tempo, pior tempo)

# Avaliação objetiva

- Erro

- This is the foundation of lossy coding techniques such as MP3, JPEG, or digital TV standards. None of the metrics presented so far reflects the impression of deviations by humans.
- O erro de uma observação pode não refletir o fiel comportamento do sistema
- Portanto, é necessário que possamos medir o erro de diversas observações do sistema
- Seja  $x$  o resultado do mundo real e  $y$  a saída do sistema, o erro seria

$$e = x - y$$

$$e = |x - y|$$

$$MAE(x, y) = \frac{1}{n} \sum_{i=1}^n |x_i - y_i|$$

# Avaliação objetiva

- Erro
  - O que temos de novidade nessa medida?

$$MSE(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2$$

$$RMSE(x, y) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2}$$

# Avaliação objetiva

- Erro

- O erro médio absoluto não é tão sensível a outliers

$$MSE(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2$$

$$RMSE(x, y) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2}$$

# Avaliação objetiva

- Erro
  - O erro médio absoluto não é tão sensível a outliers

$$\begin{aligned} PSNR(x, y) &= 10 \log_{10} \left( \frac{x_{max}^2}{MSE(x, y)} \right) \\ &= 20 \log_{10} \left( \frac{x_{max}}{RMSE(x, y)} \right) \end{aligned}$$

# Avaliação objetiva

- Perceptível?



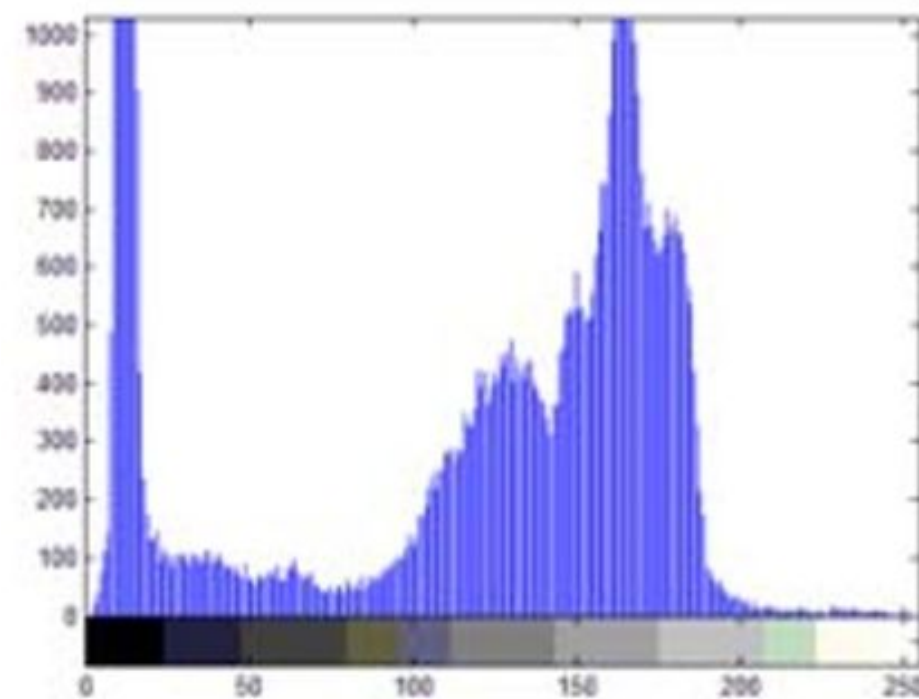
Original image (Cameraman)



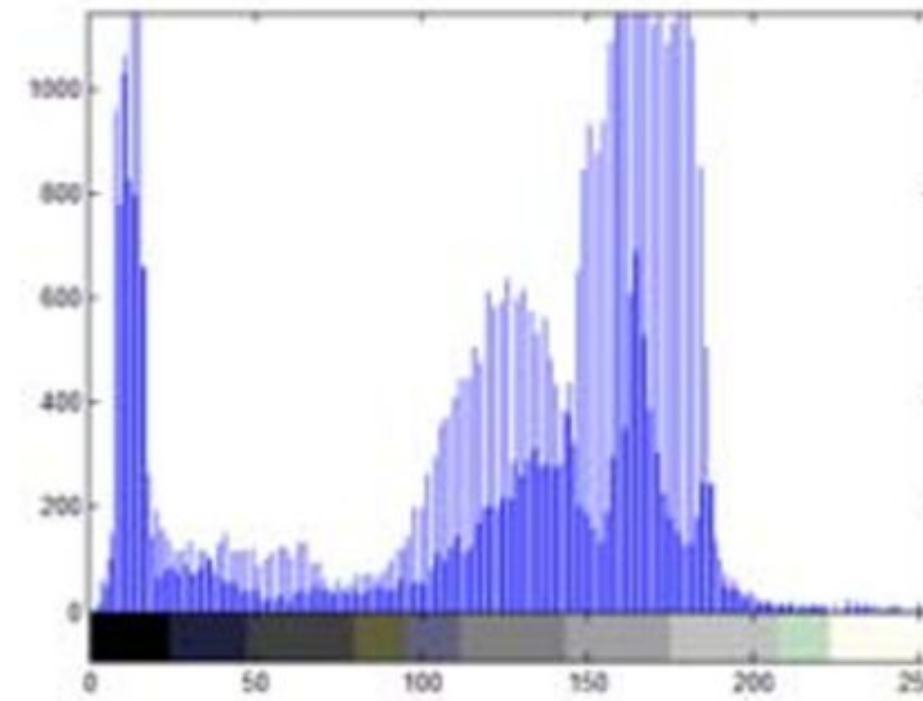
Stego image in method [8]



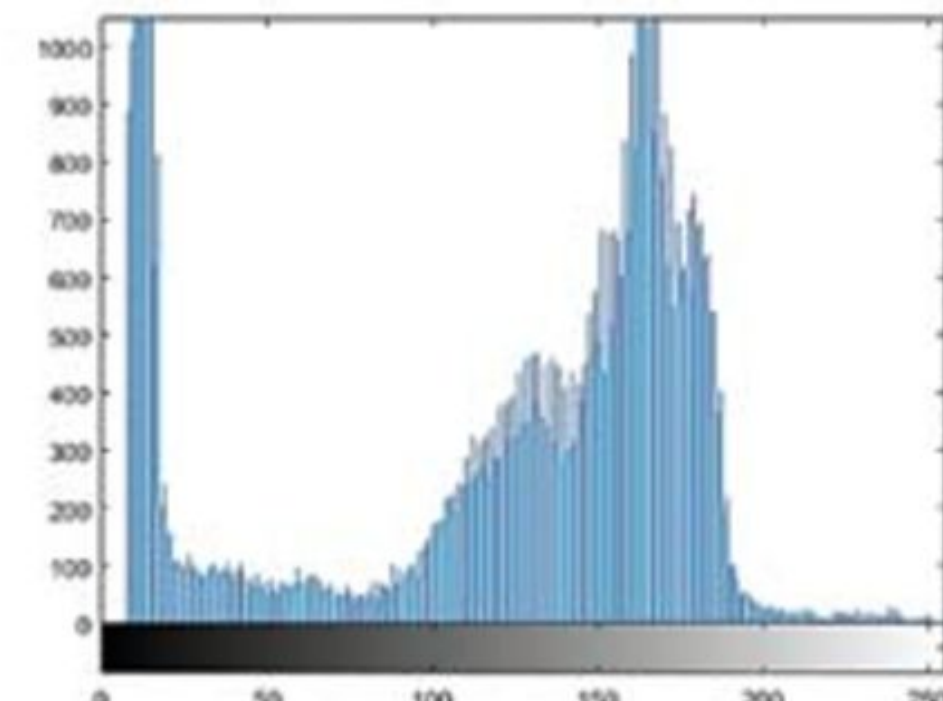
Stego image in method [38]



Original Histogram



Stego histogram in method [8]



Stego histogram in method [38]

**Fig. 1 Sample histogram of previous LSB method.** Original image (Cameraman), Stego image in method [8], Stego image in method [38], Original Histogram, Stego histogram in method [8], Stego histogram in method [38]

# Avaliação subjetiva

- Medidas que conseguem capturar mudanças em imagens, por exemplo. Essas mudanças refletiriam a avaliação humana

$$\mu_x = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\mu_y = \frac{1}{n} \sum_{i=1}^n y_i$$

$$\ell(x, y) = \frac{2\mu_x\mu_y}{\mu_x^2 + \mu_y^2}$$

# Avaliação subjetiva

- Medidas que conseguem capturar mudanças em imagens, por exemplo. Essas mudanças refletiriam a avaliação humana

$$\sigma_x = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n (x_i - \mu_x)^2}$$

$$\sigma_y = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n (y_i - \mu_y)^2}$$

$$c(x, y) = \frac{2\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2}$$

# Avaliação subjetiva

- Perceptual Evaluation of Speech Quality (PESQ)
- ITU Recommendation P.862
- Mean Opinion Score (MOS) que expressa a qualidade da voz em números desde 1.0 (worst) até 4.5 (best)
- Embora ITU recommendations varie até 5.0, PESQ tenta reproduzir os valores dados pela percepção de indivíduos que escutam ao áudio (o melhor valor estatisticamente é 4.5)

MOS	Description
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

# Avaliação subjetiva

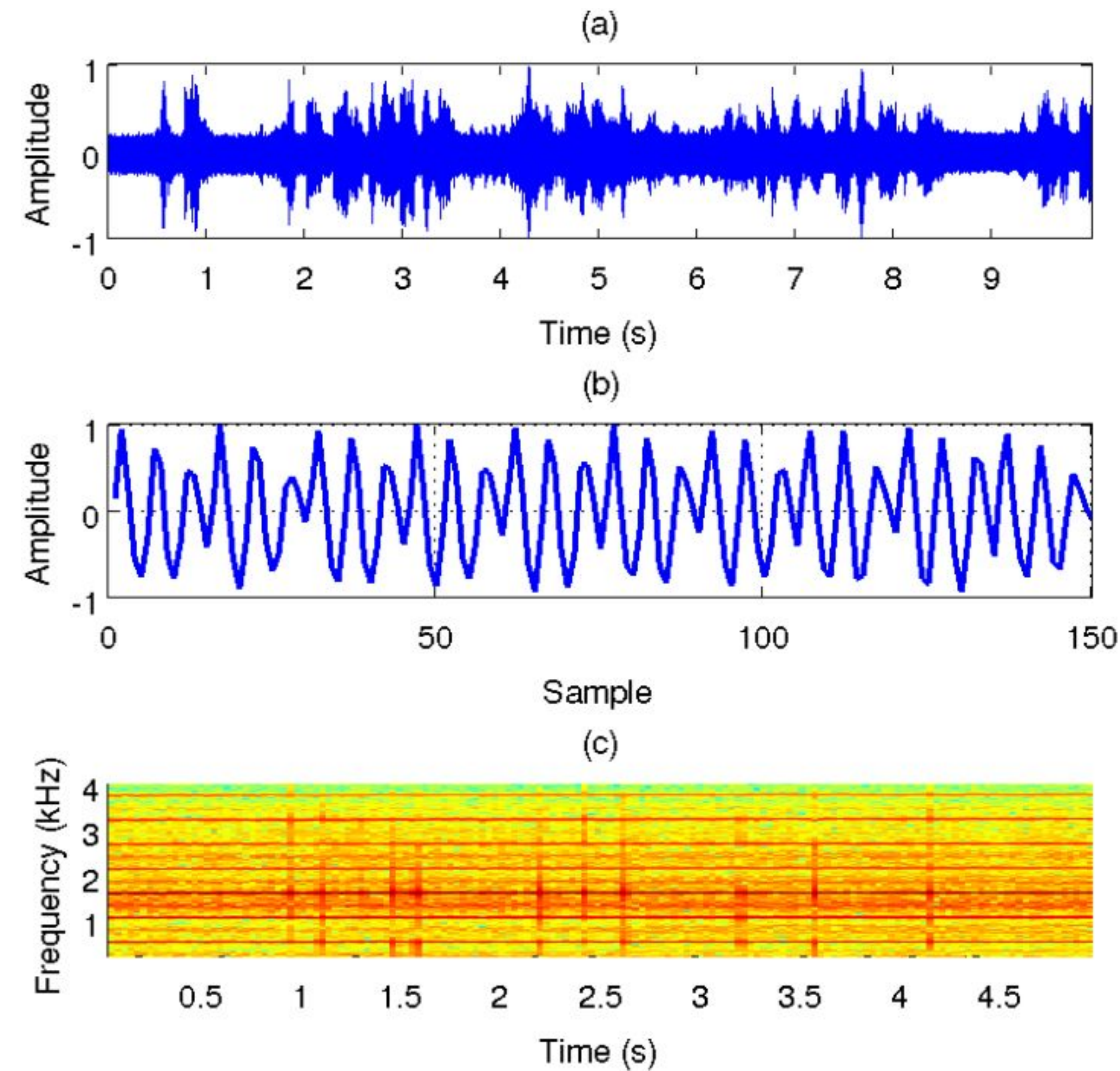
- PESQ – exemplo



MOS	Description
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

# Avaliação subjetiva

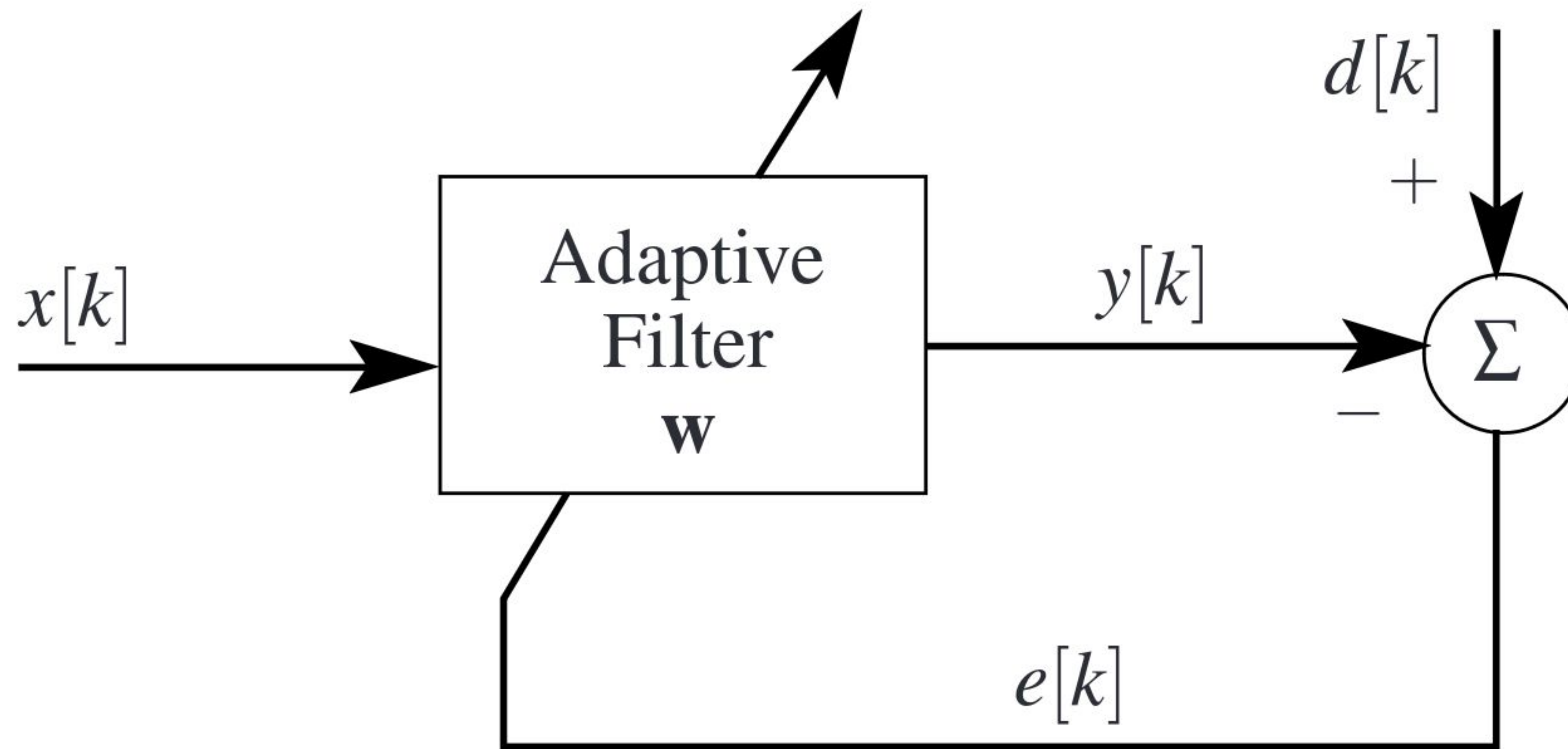
- PESQ – exemplo



MOS	Description
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

# Avaliação subjetiva

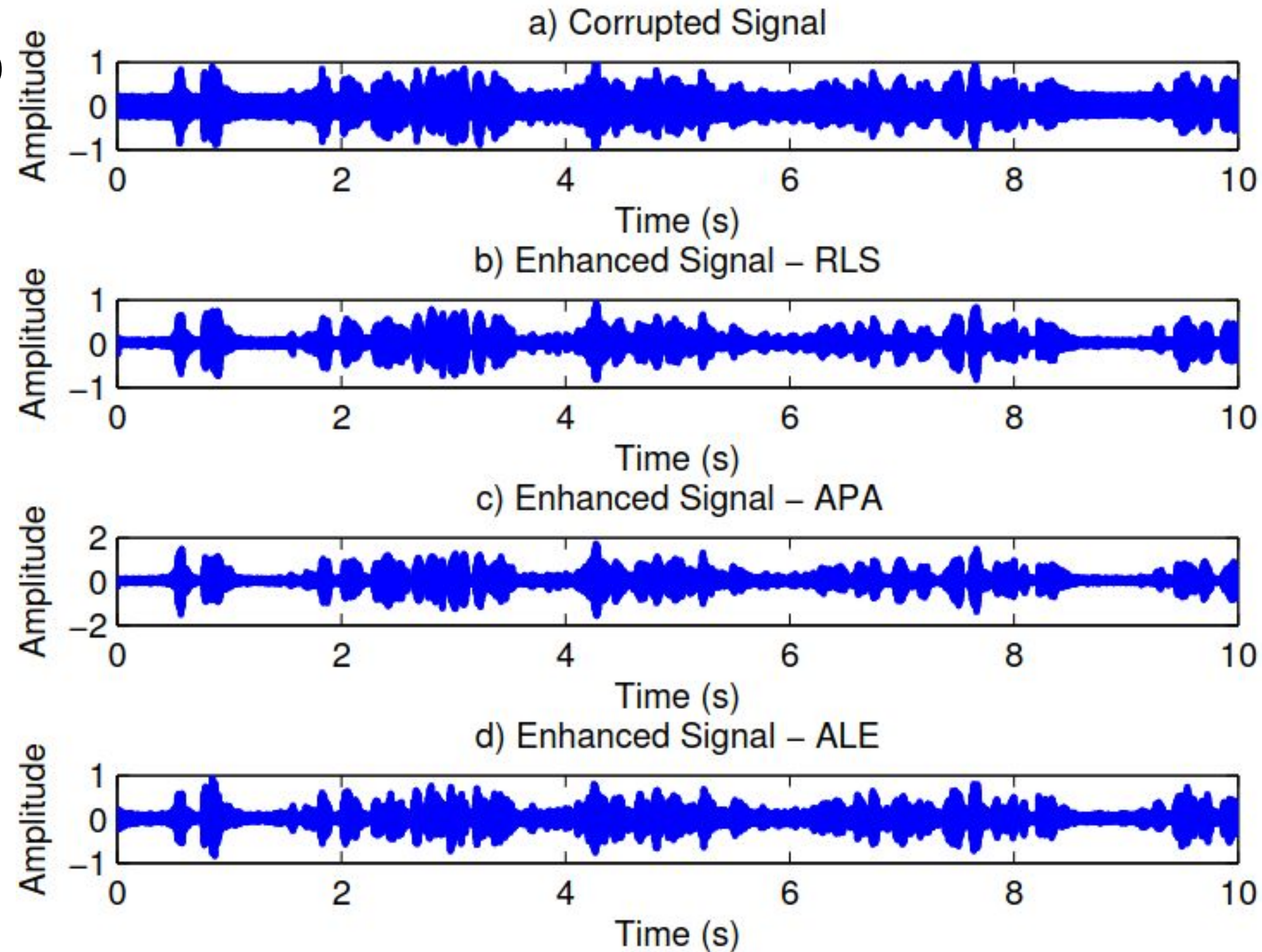
- PESQ – exemplo



MOS	Description
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

# Avaliação subjetiva

- PESQ – exemplo



MOS	Description
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

# Avaliação quantitativa

- PESQ – exemplo

#	Before filtering	ASE APA	ASE RLS	ALE APA	ALE RLS
1	1.77	<b>2.39</b>	2.24	1.85	1.94
2	2.18	2.20	<b>2.31</b>	2.07	2.08
3	1.97	1.81	<b>2.11</b>	1.97	1.93
4	2.02	1.90	<b>2.26</b>	1.96	1.90
5	1.88	2.17	<b>2.23</b>	1.99	1.92
6	1.57	1.92	<b>1.97</b>	1.72	1.74
7	1.97	2.29	<b>2.36</b>	1.97	1.92
8	1.77	2.22	<b>2.28</b>	1.91	1.86
9	1.58	<b>2.01</b>	2.00	1.68	1.56
Avg	1.86	2.10	<b>2.20</b>	1.90	1.87

MOS	Description
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

# Avaliação - análise de dados

- Sensors are typically not ideal in sense that some readouts deviate from the real values
- it may be necessary to fuse data generated by various sensors
  - True positives (TP): classificar algo como sucesso e é sucesso
  - True negative (TN): classificar algo como falha e é falha
  - False positive (FP): classificar algo como sucesso e na verdade é falha
  - False negative (FN): classificar algo como falha e na verdade é sucesso

# Avaliação - análise de dados

- Precisão  $p$

$$p = \frac{TP}{TP + FP}$$

- Recall  $r$

$$r = \frac{TP}{TP + FN}$$

- Accuracy  $acc$

$$acc = \frac{TP + TN}{TP + FP + TN + FN}$$

		Predicted	
		0	1
Actual	0	TN	FP
	1	FN	TP

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

# Avaliação - análise de dados

- Precisão  $p$

$$p = \frac{TP}{TP + FP}$$

- Recall  $r$

$$r = \frac{TP}{TP + FN}$$

- specificity

$$specificity = \frac{TN}{TN + FP}$$

		Predicted	
		0	1
Actual	0	TN	FP
	1	FN	TP

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

# Avaliação - análise de dados

- Precisão  $p$

$$p = \frac{TP}{TP + FP}$$

- Recall  $r$

$$r = \frac{TP}{TP + FN}$$

- F1 score

$$F1 = 2 \frac{p * r}{p + r}$$

		Predicted	
		0	1
Actual	0	TN	FP
	1	FN	TP

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

# Exercício proposto

- **Medir o desempenho da solução em desenvolvimento**
  - escolher inicialmente a medida de desempenho mais fácil de implementar em cada solução para obter os primeiros resultados
  - analisar os resultados (validar se as medidas objetivas estão refletindo os
- **Recalling... Definition: Validation is the process of checking whether or not a certain (possibly partial) design is appropriate for its purpose, meets all constraints, and will perform as expected**
- **Vamos, na prática entender como esses dois processos interagem no contexto do desenvolvimento de um projeto**
- **Após a escolha das medidas de desempenho que iremos empregar e de algumas validações do sistema com suporte das medidas de desempenho quantitativas (objetivas e subjetivas) podemos entrar na etapa de otimização**




IBMEC.BR

 /IBMEC

 IBMEC

 @IBMEC\_OFICIAL

 @IBMEC

 **ibmec**