

Classificação de Padrões em Escoamento Líquido-Gás Utilizando Modelo de Mistura de Gaussianas e Máquina de Vetores de Suporte

Jefferson dos Santos Ambrosio, Daniel Rodrigues Pipa, Marco José da Silva

Resumo—A classificação do padrão de escoamento é uma etapa fundamental em muitos processos que envolvem escoamento multifásico, entre eles, estimação de vazão mássica e/ou volumétrica, cálculo de fração de vazio, tamanho de bolha, fator de escorregamento etc. Neste artigo é proposta uma técnica de classificação de padrões para escoamento vertical líquido-gás. Foram utilizados dados de fração de vazio de um sensor *Wire-Mesh* para obter um histograma, que é aproximado por um modelo de mistura de gaussianas, onde os parâmetros são extraídos a partir do algoritmo *Expectation-Maximization*. Após a obtenção estes parâmetros são utilizados como entrada em um algoritmo baseado em máquina de vetores suporte multi-classes (*Multi-class Support Vector Machine* – Multiclass SVM) para enfim, fazer a identificação do padrão. Devido aos resultados, que alcançaram uma taxa de acerto entre 95 e 100% e as características físicas da análise, onde o histograma da fração de vazio independe de propriedades como condutividade e viscosidade dos fluidos, pode ser possível a aplicação dessa metodologia para a classificação de padrões em tempo real.

Palavras-Chave—Escoamento Bifásico, Aprendizado de Máquina, Algoritmo *Expectation-Maximization*.

Abstract—The classification of the flow pattern is a fundamental step in many processes that involve multiphase flow, among them, estimation of mass and/or volumetric flow, calculation of void fraction, bubble size, slip ratio, just to cite a few. In this work, a liquid-gas vertical flow pattern classification technique is propose. Void fraction data from a *Wire-Mesh* Sensor was applied to obtain a histogram, which is approximated by a Gaussian mixture model and the parameters are extracted from the *Expectation-Maximization* algorithm. After obtaining these parameters are used as input in a algorithm based on multi-class *Support Vector Machine* (multiclass SVM) to finally perform the identification of the flow pattern. Due to the results, which reached a success rate between 95 and 100%, and the physical characteristics of the analysis, where the void fraction histogram is independent of properties such as conductivity and viscosity of the fluids, it may be possible to apply this methodology to the classification of flow patterns in real time.

Keywords—Two-Phase flow, Machine Learning, *Expectation-Maximization* Algorithm.

I. INTRODUÇÃO

A identificação de padrões é um dos problemas fundamentais no estudo de escoamento multifásico. Existem várias técnicas de estimação de parâmetros nesse tipo de aplicação,

Jefferson dos S. Ambrosio, Daniel R. Pipa, Marco J. da Silva, Pós-Graduação em Engenharia Elétrica e Informática Industrial (CPGEI), Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba-PR. E-mails: ambrosioj@alunos.utfpr.edu.br, danielpipa@utfpr.edu.br, mdasilva@utfpr.edu.br. Este trabalho foi financiado pela FUNTEF-PR através do Projeto 4.01.0049 (1178) em colaboração com a Petrobras.

entre eles, tamanho de bolha, velocidades instantâneas, velocidades superficiais, razão de escorregamento etc [1]. Uma identificação correta do padrão de escoamento interfere diretamente na capacidade de estimar esses parâmetros e no desempenho dos sistemas. Isso pode ser observado em trabalhos como [2], onde primeiramente, é identificado o padrão de escoamento utilizando uma técnica denominada *Fisher Discriminant Analysis* (FDA) e em seguida é escolhido um modelo de cálculo de fração de vazio referente a cada padrão. Esses modelos são obtidos por um algoritmo baseado em máquina de vetores suporte (*Support Vector Machine* – SVM). Isso deixa claro que uma identificação errada leva a uma escolha de modelo errada e conseqüentemente, falhas de estimação e redução de desempenho.

Existem diversas abordagens quando se trata o problema da identificação. Em [3] foi feito um estudo de como o histograma ou PDF (*Probability density Function*) do sinal de fração de vazio se relaciona com o padrão de escoamento. A referência [4] combina este estudo com técnicas de aprendizado de máquina (*Machine Learning*) utilizando redes neurais (*Artificial Neural Networks* – ANNs) e algoritmos de clusterização como: *self-organized map*, *K-means* e *Fuzzy C-means*. Em [5] também é discutida uma análise utilizando ANNs, porém, é feito um pré-processamento através de uma técnica denominada *Natural Logarithmic Normalization*, uma transformação não linear que traz uma maior clareza e separabilidade aos dados que alimentarão a rede neural. Voltando à análise do histograma da fração de vazio, em [6] são usados sensores de fibra óptica para obter um sinal de fração de vazio. A partir desse sinal são calculados quatro momentos estatísticos: média, desvio padrão, assimetria e curtose. Então esses dados são usados para alimentar um algoritmo baseado em SVM para a identificação do padrão. Seguindo essa tendência, [7] faz um levantamento de vários trabalhos que empregam diversas topologias de sensores, técnicas de *machine learning* e técnicas de processamento de sinais.

No presente trabalho serão utilizados dados de fração de vazio oriundos de um sensor *Wire-Mesh*. Esse sensor é muito utilizado em problemas de escoamento multifásico, como pode ser visto em [8] e [9]. A partir da série temporal dos dados de fração de vazio é obtido um histograma que em seguida é aproximado por um modelo de mistura de gaussianas gerando os parâmetros que serão utilizados em um algoritmo de classificação baseado no Multiclass SVM. Nos dados experimentais foram observados três padrões de escoamento: 'bolhas' (*Bubble flow*), 'agitado' (*Churn flow*) e

'golfadas' (*Slug flow*).

II. ESCOAMENTO BIFÁSICO VERTICAL E FRAÇÃO DE VAZIO

Vários fatores influenciam na definição do padrão de escoamento. Entre eles, a inclinação da tubulação, o tipo de fluido, características físicas destes fluido etc. No caso da inclinação, que nos casos mais divergentes, que são o horizontal e o vertical, existem alguns padrões que se mantem, e alguns padrões distintos que aparecem em cada caso. Em relação aos fluidos, considerando um escoamento bifásico, podemos ter uma interface líquido-líquido, líquido-gás, gás-líquido e assim por diante. Algo que também pode ocorrer é um escoamento bifásico do mesmo fluido, porém em estados físicos diferentes, como é o caso do escoamento de água e vapor na mesma tubulação. A Figura 1 indica um tipo de classificação de padrões para escoamento vertical do tipo líquido-gás definida em [10]. Com o aumento da velocidade relativa entre gás e líquido, podemos ver como a interface é modificada no escoamento. Em baixas velocidades de gás verifica-se a estrutura de bolhas dispersas na fase líquida. Com o aumento dessa velocidade acontece uma coalescência entre essas pequenas bolhas gerando estruturas de gás significativamente maiores, denominadas bolhas de Taylor [10]. A partir do momento em que essa velocidade aumenta ainda mais, é possível observar um comportamento caótico com uma maior presença de gás e em seguida, com um novo aumento da velocidade relativa, forma-se uma coluna central de gás que contém algumas gotículas de líquido.

Poucas aplicações possuem tubulações transparentes que permitem uma visualização do desenvolvimento do escoamento como mostra a Figura 1, por isso, são necessárias outras ferramentas para que a classificação do padrão seja feita. Uma delas, proposta ainda em [10], define um mapa desses padrões através da velocidade superficial de cada fase e do diâmetro da tubulação, como pode ser visto na Figura 2.

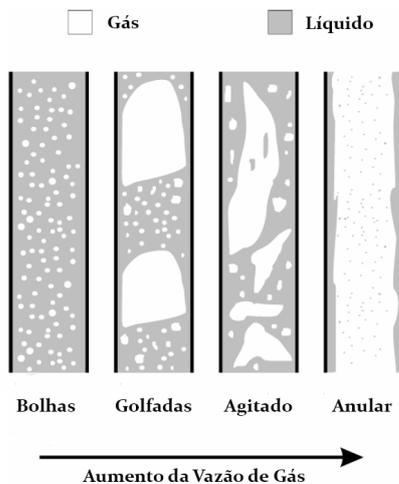


Fig. 1. Padrões típicos de escoamento bifásico líquido-gás. Adaptado de [9].

Nos estudos a seguir foram utilizados dados de escoamento vertical água-ar gerados em uma planta de teste vinculada

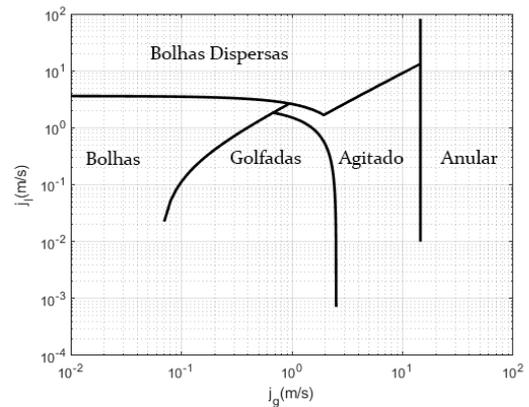


Fig. 2. Relação do padrão de escoamento com as velocidades superficiais como definido em [10].

ao instituto Helmholtz-Zentrum Dresden-Rossendorf (HZDR) utilizando um sensor *Wire-Mesh* como pode ser verificado no relatório técnico [11]. Devido ao projeto "Construction and execution of experiments at the multi-purpose thermal hydraulic test facility TOPFLOW for generic investigations of two-phase flows and the development and validation of CFD codes", vários experimentos de escoamento bifásicos foram realizados. Neste trabalho foram utilizados dados de 67 pontos de operação de escoamento vertical acendente em tubos de duas polegadas. Todas as informações de estrutura e metodologia das medições, além de uma diagrama da planta de teste podem ser encontradas no relatório técnico já citado. O processamento dos dados e a implementação dos algoritmos foram feitas através do *software* MATLAB[®].

A Figura 3 apresenta um esquemático simplificado do sensor *Wire-Mesh* que pode ser encontrado em [9]. Através das intersecções de condutores é possível identificar bolhas e aproximar seus tamanhos. A partir do processamento desses dados temos o sinal de fração de vazão na seção transversal do tubo. Uma análise mais aprofundada desse sensor é feita em [8], [9] e [12]. O sensor utilizado forma uma matriz de condutores de 16x16 e diâmetro interno de 51.2 mm. Foi utilizada uma frequência de amostragem de 2,5 kHz e um período de amostragem de 10 segundos. A partir do sinal de série temporal da fração de vazão é calculado um histograma, seguindo a mesma abordagem de [3]. Um exemplo desses dados pode ser visualizados na Figura 4.

Foram utilizadas 67 medições, onde cada ponto de operação é definido por um par de velocidades superficiais de líquido e gás, obtidos por sensores de referência. A Figura 5 mostra os dados do HZDR classificados de acordo com o mapa definido em [10].

III. MODELO DE MISTURA DE GAUSSIANAS E MÁQUINA DE VETORES SUPORTE NA CLASSIFICAÇÃO DE PADRÕES DE ESCOAMENTO

Uma vez com os dados dos histogramas da fração de vazão é necessário extrair informações para efetuar a identificação do padrão. O método proposto nesse trabalho faz uma

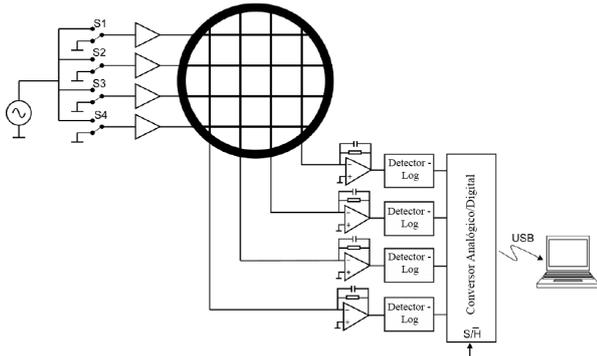
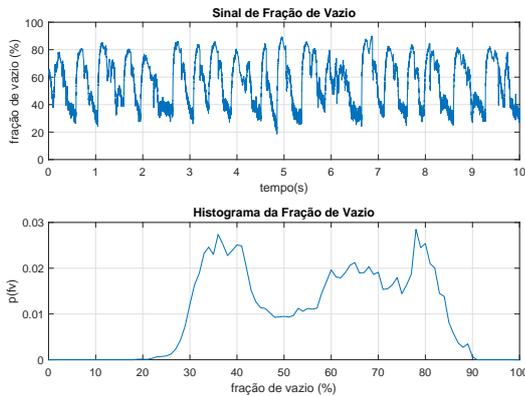

 Fig. 3. Esquemático simplificado de um sensor *Wire-Mesh*. Adaptado de [9].


Fig. 4. Sinal de fração de vazio e seu respectivo histograma.

aproximação deste histograma através de um modelo de mistura de gaussianas. Sendo $f(x)$ uma PDF gaussiana

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (1)$$

onde μ e σ são a média e o desvio padrão da distribuição respectivamente. Podemos escrever o modelo de mistura de gaussianas como

$$g(x) = \sum_{j=1}^K \phi_j f(x; \mu_j, \sigma_j), \quad (2)$$

sendo K o número de componentes do modelo e ϕ_j um fator de proporção para cada componente tal que $\sum_{j=1}^K \phi_j = 1$. Tais modelos podem ser usados tanto para a aproximação de PDFs quanto para modelos de clusterização.

Como as PDFs obtidas a partir do sinal de fração de vazio de escoamentos verticais apresentavam um comportamento uni, bi ou tri-modal. Foi definido um modelo de mistura de três componentes, gerando um vetor de parâmetros $\theta_i = (\phi_{1i}, \phi_{2i}, \phi_{3i}, \mu_{1i}, \mu_{2i}, \mu_{3i}, \sigma_{1i}, \sigma_{2i}, \sigma_{3i})$. Sendo $i = 1, \dots, 67$. Ou seja, um vetor de parâmetros para cada ponto de operação. Para estimar esse vetor de parâmetros, foi utilizado um algoritmo chamado *Expectation-Maximization* (EM). Este algoritmo foi formalizado em [13] como um estimador de máxima verossimilhança (*Maximum Likelihood Estimator* – MLE) e é bastante difundido na estimação de parâmetros em

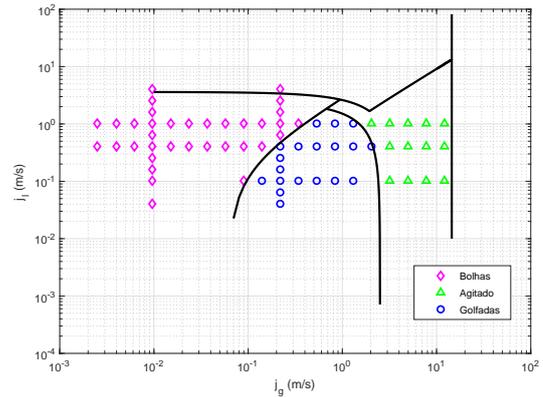


Fig. 5. Dados do HZDR classificados de acordo com o mapa definido em [10].

modelos de mistura e problemas de clusterização, como pode ser visto em [14], [15] e muitos outros trabalhos nas áreas de processamento estatístico de sinais e teoria de estimação.

Com isso, para cada ponto de operação, foi executado o algoritmo EM para a estimação do vetor de parâmetros θ_i e aproximação do histograma pelo modelo de mistura definido na Equação 2. A Figura 6 mostra um exemplo para um dos pontos de operação.

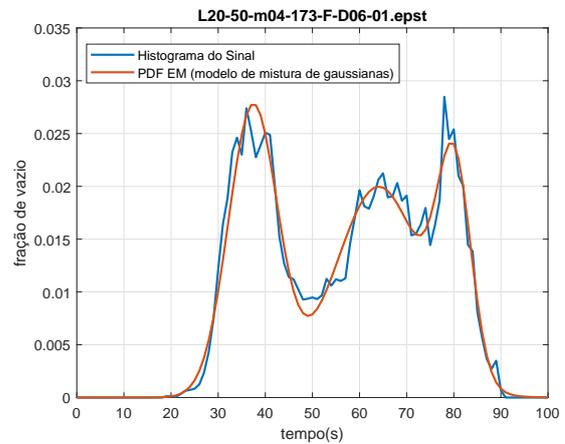


Fig. 6. Histograma da fração de vazio e PDF gerada pelo modelo de mistura de gaussianas através dos parâmetros obtidos pelo algoritmo EM.

Na última etapa do processo, os parâmetros θ_i gerados pelo EM, para cada ponto de operação são utilizados como variáveis de entrada em um algoritmo baseado no Multi-class SVM [16] implementado através da função "fitcecoc" do MATLAB[®], que implementa um classificador multi-classe a partir da abordagem "one-against-one" [17]. Com isso, a classificação do padrão de escoamento é realizada, como pode ser visto no fluxograma da Figura 7.

Dos 67 pontos de operação, 35 pontos pertencem ao padrão 'bolhas', 19 pontos ao padrão 'golfadas' e 13 pontos ao padrão 'agitado'. Para a verificação do algoritmo foram utilizados aproximadamente 70% dos pontos de operação para treino do Multiclass SVM e 30% dos pontos foram usados para teste. Ou seja, dos 35 pontos do padrão 'bolhas', 25 foram usados para

treino e 10 para teste. Dos 19 pontos do padrão 'golfadas', 13 foram utilizados para treino e 6 para teste. E por fim, dos 13 pontos do padrão 'agitado', 9 foram usados para treino e 4 para teste.

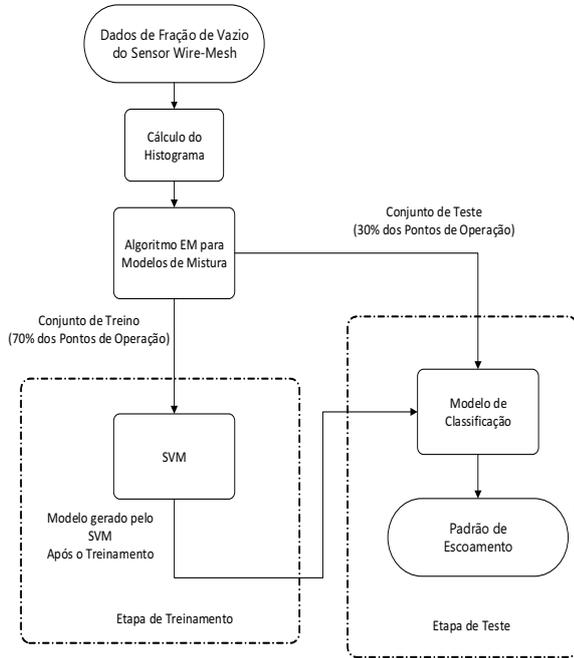


Fig. 7. Fluxograma do processo de classificação de padrão.

IV. RESULTADOS E DISCUSSÃO

Após alguns testes preliminares, foi possível verificar que tanto as médias quando os fatores de proporção do modelo de mistura tinham uma correlação com o padrão de escoamento. Mas o mesmo não acontecia com as informações de desvio padrão. Por isso foram realizados dois testes: o primeiro utilização um vetor de parâmetros com a informação de variância, $\theta_i = (\phi_{1i}, \phi_{2i}, \phi_{3i}, \mu_{1i}, \mu_{2i}, \mu_{3i}, \sigma_{1i}, \sigma_{2i}, \sigma_{3i})$ e um segundo teste onde o vetor de parâmetros desconsiderava os desvios padrões, $\hat{\theta}_i = (\phi_{1i}, \phi_{2i}, \phi_{3i}, \mu_{1i}, \mu_{2i}, \mu_{3i})$. Na Figura 8 podemos ver a relação entre os histogramas da série temporal de fração de vazio e cada padrão de escoamento, como discutido em [18]. Conforme o padrão de escoamento muda, é possível observar uma alteração significativa no posicionamento dos picos do histograma e as diferenças de amplitude. Isso reforça a ideia anterior de que as médias e os fatores de proporção estão diretamente relacionados ao padrão de escoamento. No caso do desvio padrão, acabam havendo sobreposições, o que dificulta a separabilidade dos dados.

Na Figura 9 temos o resultado do Multiclass SVM que utilizou o vetor de parâmetros θ_i que continha a informação dos desvios padrões do modelo de mistura. A legenda mostra que os padrões de escoamento foram diferenciados por cor e os pontos de treinamento e teste foram diferenciados pelo marcador. Podemos ver um ponto indicado em vermelho. Esse ponto mostra um erro de classificação de acordo com o mapa definido em [10]. O mesmo ponto é classificado como padrão 'Bolhas' pelo mapa. Porém, foi classificado como padrão 'Golfadas' pelo algoritmo de classificação.

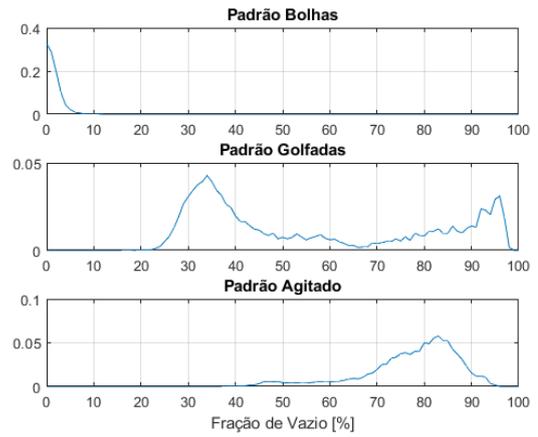


Fig. 8. Histogramas de escoamento líquido-gás.

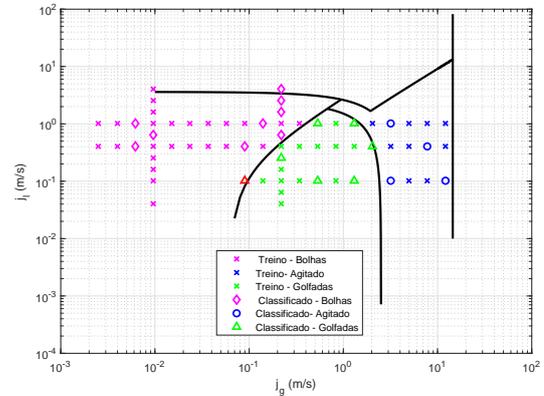


Fig. 9. Resultados do treinamento considerando os desvios padrões do modelo de mistura como variáveis de entrada.

Um ponto interessante a ser observado é que o único erro de classificação do algoritmo se encontra próximo a região de fronteira no mapa, porém, essas regiões são chamadas de regiões de transição. E em muito casos, a classificação perto dessas regiões é questionável ou ambígua, pois o escoamento pode apresentar características de ambos os padrões próximos a essa região.

O mapa de resultados mostra cada ponto de acordo com suas velocidades superficiais, mas essas velocidades são desconhecidas a priori. Essas informações só estão disponíveis pois foram utilizados sensores de referência para obtenção dos mesmos. No contexto deste trabalho, os únicos dados disponíveis são os de fração de vazio, todas as informações dos mapas relacionados às velocidades superficiais são para comparação e validação da abordagem estudada. Como algoritmo de classificação usa apenas a informação da fração de vazio, um único erro próximo de uma zona de fronteira acaba sendo uma boa característica do algoritmo, pois como dito anteriormente, as classificações de padrões de escoamento nessa região tendem a ser questionáveis.

Na Figura 10 podemos ver um desempenho ainda melhor. Neste caso, o algoritmo utilizou o vetor de parâmetros $\hat{\theta}_i$, que não considerava as informações dos desvios padrões do

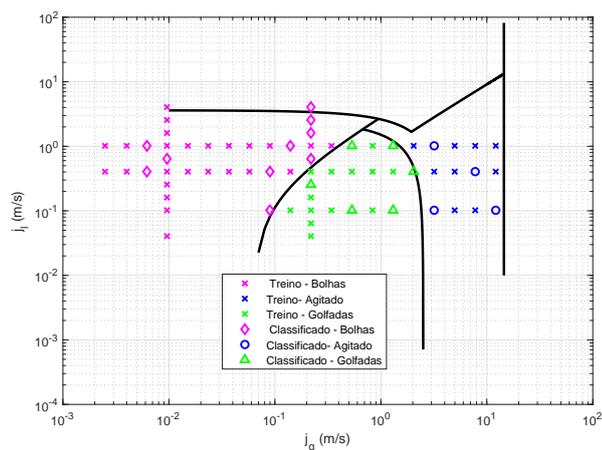


Fig. 10. Resultados do treinamento desconsiderando os desvios padrões do modelo de mistura como variáveis de entrada.

modelo de mistura. O aumento do desempenho corrobora com a ideia de que a informação do desvio padrão tem uma menor contribuição em relação à classificação do padrão com a abordagem utilizada.

V. CONCLUSÃO

Nesse artigo é discutida uma abordagem para classificação de padrões de escoamento líquido-gás em escoamento vertical. Essa abordagem consiste em utilizar os dados de fração de vazio obtidos por um sensor *Wire-Mesh*, calcular o histograma e utilizar o algoritmo EM para aproximá-lo por um modelo de mistura de gaussianas. Em seguida, os parâmetros calculados pelo EM são utilizados como variáveis de entrada no algoritmo Multiclass SVM para que a classificação do padrão seja feita.

A partir dos resultados apresentados, foi possível verificar que o caso onde a informação do desvio padrão foi utilizado, ocorreu apenas um erro. Como os testes foram realizados em 20 pontos de operação, isso resulta em uma taxa de acerto de 95%. No caso onde o Multiclass SVM não utiliza a informação do desvio padrão, essa taxa de acerto subiu para 100%.

Com a eficiência desse algoritmo, podemos usá-lo em outros modelos ou algoritmos para outras aplicações. Por exemplo, para estimação de vazão, alguns modelos fazem estimativas indiretas e dependem do padrão de escoamento. Outra característica desse algoritmo, é que provavelmente, a relação entre o histograma e os padrões de escoamento se mantém mesmo com algumas mudanças de diâmetro de tubos e características do fluido. Porém, se essa possibilidade for confirmada, a fase de treinamento e calibração do algoritmo poderá ser executada *offline*. O que eliminaria um dos grandes vieses dessa abordagem de treinamento em sistemas industriais.

Os resultados alcançados abrem a possibilidade para mais testes e mais estudos. Por exemplo, pode ser feita uma análise para otimizar a escolha dos pontos que serão utilizados para treinamento e qual seria a quantidade ideal para cada padrão. Também é possível fazer um estudo de como essa abordagem funciona em escoamentos horizontais e inclinados.

AGRADECIMENTOS

Os Autores agradecem a Petrobras pelo apoio e financiamento do trabalho.

REFERÊNCIAS

- [1] G. Falcone, G. Hewitt, and C. Alimonti, *Multiphase flow metering: principles and applications*. Elsevier, 2009, vol. 54.
- [2] H. Li, H. Ji, Z. Huang, B. Wang, H. Li, and G. Wu, "A new void fraction measurement method for gas-liquid two-phase flow in small channels," *Sensors*, vol. 16, no. 2, p. 159, 2016.
- [3] G. Costigan and P. Whalley, "Slug flow regime identification from dynamic void fraction measurements in vertical air-water flows," *International Journal of Multiphase Flow*, vol. 23, no. 2, pp. 263–282, 1997.
- [4] E. Rosa, R. Salgado, T. Ohishi, and N. Mastelari, "Performance comparison of artificial neural networks and expert systems applied to flow pattern identification in vertical ascendant gas-liquid flows," *International Journal of Multiphase Flow*, vol. 36, no. 9, pp. 738 – 754, 2010. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0301932210000911>
- [5] M. Al-Naser, M. Elshafei, and A. Al-Sarkhi, "Artificial neural network application for multiphase flow patterns detection: A new approach," *Journal of Petroleum Science and Engineering*, vol. 145, pp. 548 – 564, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0920410516302479>
- [6] I. Benito and N. W. Mureithi, "Identification of two-phase flow patterns using support vector classification," in *ASME 2017 Pressure Vessels and Piping Conference*. American Society of Mechanical Engineers, 2017, pp. V004T04A027–V004T04A027.
- [7] Y. Yan, L. Wang, T. Wang, X. Wang, Y. Hu, and Q. Duan, "Application of soft computing techniques to multiphase flow measurement: A review," *Flow Measurement and Instrumentation*, vol. 60, pp. 30–43, February 2018. [Online]. Available: <http://kar.kent.ac.uk/66012/>
- [8] R. Kipping, R. Brito, E. Scheicher, and U. Hampel, "Developments for the application of the wire-mesh sensor in industries," *International Journal of Multiphase Flow*, vol. 85, pp. 86 – 95, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0301932215302019>
- [9] M. J. Da Silva, "Impedance sensors for fast multiphase flow measurement and imaging," Ph.D. dissertation, Technische Universität Dresden, 2008.
- [10] Y. Taitel, D. Bornea, and A. E. Dukler, "Modelling flow pattern transitions for steady upward gas-liquid flow in vertical tubes," *AICHE Journal*, vol. 26, no. 3, pp. 345–354, 1980. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/aic.690260304>
- [11] H. Schütz, P. Pietruske, A. Manera, H. Carl, M. Beyer, and H.-M. Prasser, "Experiments on upwards gas/liquid flow in vertical pipes," *Forschungszentrum Dresden-Rossendorf e.V. Institut für Sicherheitsforschung*, Tech. Rep., September 2007.
- [12] H.-M. Prasser, A. Böttger, and J. Zschau, "A new electrode-mesh tomograph for gas-liquid flows," *Flow Measurement and Instrumentation*, vol. 9, no. 2, pp. 111 – 119, 1998. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0955598698000156>
- [13] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.
- [14] G. McLachlan and T. Krishnan, *The EM algorithm and extensions*, 2nd ed., ser. Wiley series in probability and statistics. Hoboken, NJ: Wiley, 2008.
- [15] M. Zanetti, F. Bovolo, and L. Bruzzone, "Rayleigh-rice mixture parameter estimation via em algorithm for change detection in multispectral images," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5004–5016, 2015.
- [16] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2013.
- [17] J. Milgram, M. Cheriet, and R. Sabourin, "'one against one'" or "'one against all'": Which one is better for handwriting recognition with SVMs?" in *tenth international workshop on Frontiers in handwriting recognition*. SuviSoft, 2006.
- [18] H. Shaban and S. Tavoularis, "The wire-mesh sensor as a two-phase flow meter," *Measurement Science and Technology*, vol. 26, no. 1, p. 015306, 2014.