

On the pattern recognition of Itô processes in market price data

S. N. Onyango¹ & M. Ingleby²

¹*Faculty of Science, Department of Mathematics and Applied Statistics, Maseno University, Kenya*

²*School of Computing and Engineering, University of Huddersfield, UK*

Abstract

We introduce a highly error resistant method of extracting Itô processes as applied to market data. This method is inspired by an AI method known as Hough transforms (HT). The HT method has been used in extracting geometric shape patterns from noisy and corrupted image data. We use this method to extract simultaneously geometric Brownian motion trends and market parameters (volatility and mean) from simulated price histories and real-market price data. It turns out that this approach is an effective method of extracting market parameters and market processes for both simulated and real-world market price data.

Keywords: Itô processes, geometric Brownian motion, artificial intelligence (AI), Hough transforms, price histories, market processes, pattern recognition.

1 Introduction

Artificial Intelligence (AI) is the application of science to the development of systems that appear to manifest intelligence of the human or mammalian sort. It is a research field where computer science intersects with other fields such as, philosophy, psychology, linguistics, finance and economics among others.

The characteristics of a successful recogniser of patterns corrupted by noise are robustness of the following different types:

- resistance to statistical errors in the capturing process ('Gaussian noise'),
- resistance to extreme errors ('impulsive noise'),
- resistance to loss of data ('pattern occlusion').

There is a growing literature on attempts to apply multi-layer artificial neural nets (ANNs) to financial market data (Kingdon [8]; Zapranis and Refenes [16]).



In this paper, however, we apply a method that has been less publicised than the ANN: the *Hough transform*.

1.1 Pattern recognition in asset price history

The primary aim of pattern recognition is to identify objects and images from their shapes, forms, outlines, colour, surface extraction, temperature or sound or any sensed attributes- usually by automatic means (Weik [15]). Pattern recognition has developed into a powerful tool in many disciplines. One of the virtues of pattern recognition is the ability to simplify complex environments by picking out pattern from noise. Another virtue is its reliance on algorithms that are executable with tremendous simplicity, clarity, speed, and power. Pattern recognition supports effective and rapid decision-making in any business environment, however time sensitive it may be. In an increasingly changing, chaotic, and complex business environment, understanding and internalising market patterns is vital and crucial, for helping management and other market players to make safe, rapid, and effectively supported decisions.

In financial markets, it is an established notion that prices of underlying assets move in patterns and that these patterns can be used to forecast asset prices statistically. Every financial analyst needs to define and search out meaningful patterns from historical time series quickly and efficiently. Such discovery of pattern can be time consuming and technically challenging, especially in the absence of high-level pattern recognition techniques.

In this paper, our preferred technique of pattern recognition is inspired by the classical Hough transform (HT). We give a summary of the Hough transform in the next section and adapt it to identifying patterns in financial markets that is to estimate robustly the parameters that characterise the stochastic dynamics of changing asset prices.

2 The Hough transforms (HT)

The Hough transform (HT) (Hough [7], Duda and Hart [5]; Leavers [9]; Princen et al. [13]; Toft [14]) is a standard tool for extraction of geometrical primitives such as line-segments and arcs from noisy digital images. Hough transforms change the mode of presentation of data set in order to ease detection of a specific geometric form being sought.

The characteristic relation (1) below, of the sought-for feature is back-projected in the space of pattern parameters, $(\alpha_1, \alpha_2, \dots, \alpha_n)$. A pixel (x, y) that lies on a parametric curve of the characteristic relation that is inconsistent may be represented in parametric space, if

$$f((x, y), (\alpha_s)) = 0 \quad (1)$$

holds, then the idea of the Hough transform is to convert a pixel position (x, y) into a relation between pattern parameters by fixing (x, y) in (1) above.

One of the main advantages of the Hough transforms is its robustness to noise and occlusion. This is due to the fact that each image point is considered



independently of the other points. Hence, within reason, the occlusion of valid pixels only alters peak intensities in parameter space (McDonald 2001). Similarly, the addition of a few noise pixels only adds low-level peaks in parameter space. In financial markets, a direct consequence is the capability of transforms to detect broken line patterns as may be reflected in stock prices during nights, or weekends when trading is suspended to next day or to Monday of next week (Onyango [12]). Hough transforms are also capable of detecting several different features in a given image in one pass. That is, it can simultaneously accumulate ‘votes’ for several peaks in the accumulator array, if the given pattern consists of several primitives, as in the case of several processes (Onyango, op cit).

2.1 The (ρ-θ) Hough transforms

The slope-intercept form of a line, $(y = mx + c)$, described above has a problem with the vertical lines, both m (gradient) and c (y-intercept) are infinite i.e. they are unbounded. This problem is eliminated by expressing a line passing through point (x, y) in standard polar $(\rho-\theta)$ form instead of the Cartesian (m, c) form. That is

$$x \cos \theta + y \sin \theta = \rho, \tag{2}$$

where ρ is the perpendicular distance to the line AB from the origin and θ is the angle between this perpendicular or ‘normal’ and the x-axis. By convention $0 \leq \theta < \pi$.

If the point (x, y) is fixed, equation (2) describes a sinusoidal curve in (θ, ρ) parametric space. A pixel at point (x_i, y_i) in image votes into cells through which the sinusoidal curve (2) passes.

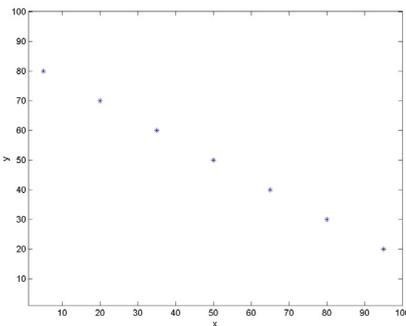


Figure 1: An image of collinear points.

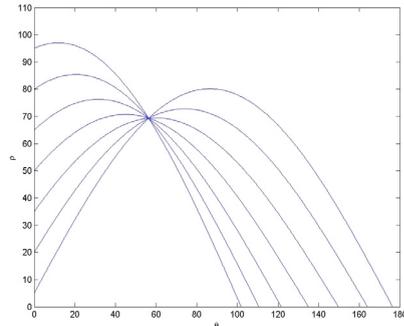


Figure 2: The Hough transforms for collinear points.

We illustrate the Hough transforms in parametric $(\rho-\theta)$ by considering a set of 7 collinear image points in Fig. 1. The corresponding sinusoidal voting curves intersect at a single point in parameter space as depicted in Fig.2. Fig.3 depicts a 3-dimensional accumulator array with the peak clearly labelled. Since HT is error resistance, the transform has been adapted to estimating the parameters of

deterministic dynamical laws that model the evolution of measurable outputs from engineering equipment (Aguado et al. [1]; Flint et al. [6]).

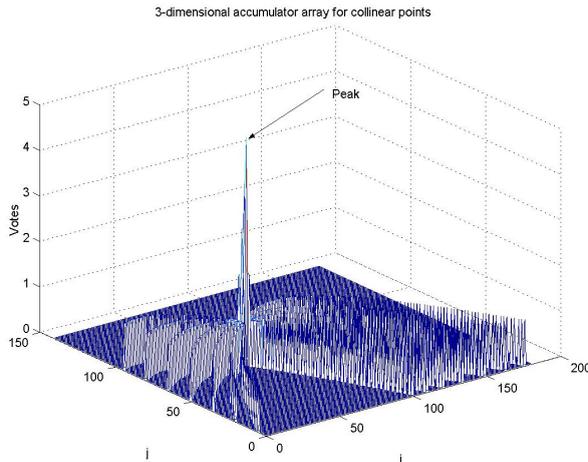


Figure 3: A 3-dimensional accumulator array (butterfly shape) for the collinear points of Fig. 1.

2.2 Detection of two processes

Hough transform can be used to detect simultaneously two or more patterns (or processes in a given dynamical system). To illustrate this property, two lines of collinear points are used (Figure 4) and the corresponding Hough transform is shown in Figure 5 and Figure 6 depicts a 3-dimensional accumulator array for the two processes of Figure 4. Two peaks clearly labelled depict the two processes.

In finance markets it is usually assumed during analysis that there is only one stochastic process, the linear geometric Brownian motion, driving the stock prices and hence the prices of derivatives. This might not be true in some markets. There might be more than one process within the time period of trade. Nevertheless, a generalisation of the Hough transforms is capable of detecting such stochastic processes in a financial market.

We recall that Figure 1 shows that a single line defines a well-delineated peak as shown in Figure 2 and Figure 3. Several lines will show several peaks in the accumulator. In particular, two lines of collinear points, representing two processes (see Figure 4), will show two peaks in the accumulator array as depicted in Figure 6. If the number of processes is not known in advance, one can discover how many processes were at work generating the data by simply counting the peaks in the accumulator.

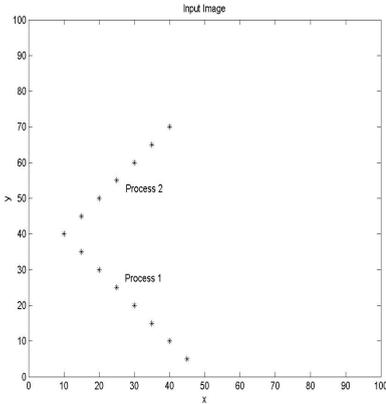


Figure 4: Images of two simple processes.

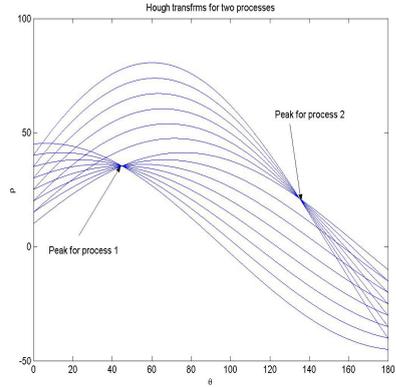


Figure 5: The corresponding Hough transforms of two processes.

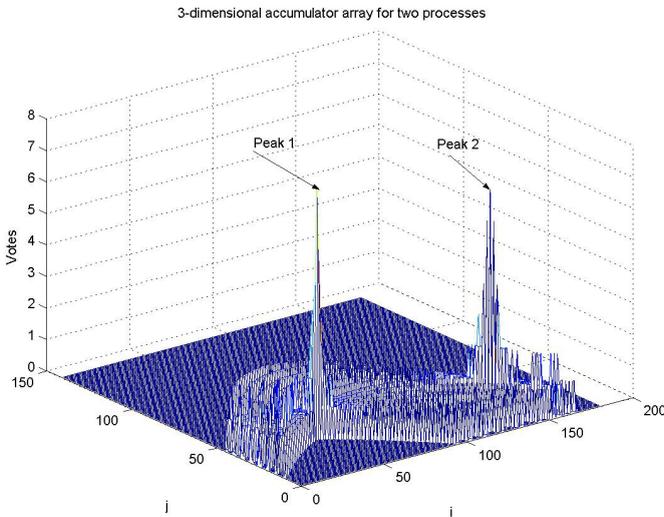


Figure 6: 3-dimensional accumulator array for two processes indicated by two peaks.

3 Adapting Hough transform to geometric Brownian motion

In this section we adapt the Hough transforms to simulated geometric Brownian motion. We introduce and use the grey-scale voting techniques to calculate the area of cell in the accumulator array. This technique is then used to test geometric Brownian motion.

3.1 Application to quadratic volatility function

One of the assumptions of Black-Scholes-Merton model (Black and Scholes [3]; Merton [11]) is that the price of an underlying asset follows the geometric Brownian motion given as

$$dS(t) = \mu S(t)dt + \sigma S(t)dZ \tag{3}$$

where S is the price of the underlying, μ is the rate of increase, σ is the volatility of the underlying and $dZ \sim N(0,dt)$. The solution of equation (3) using Itô's lemma is given as

$$\log S(t) = \log S(t_0) + (\mu - \sigma^2 / 2)(t - t_0) + \sigma(Z(t) - Z(0)) \tag{4}$$

This can also be expressed as

$$\log(S(t) / S(t_0)) = \mu^* (t - t_0) + \sigma \varepsilon \sqrt{(t - t_0)} \tag{5}$$

where $\mu^* = \mu - \sigma^2/2$ and $Z(t) - Z(t_0) = \varepsilon \sqrt{(t - t_0)}$.

From (5) we have

$$\sigma^2 = f(\mu_j, \sigma) = \frac{1}{n-2} \left[\sum_{i=1}^n \left(\frac{\log(S(t_i) / S(t_{i-1}))}{\sqrt{(t_i - t_{i-1})}} \right)^2 - 2\mu \sum \log(S(t_i) / S(t_{i-1})) + \mu^2 \sum (t_i - t_{i-1}) \right] \tag{6}$$

where $\sum \varepsilon^2 \sim \chi^2_{(n-2)}(n-2, 2n-4)$.

We adapt the Hough transforms in (2) to a stochastic dynamics by replacing pixel coordinates (x, y) with a sequence of $(\log(S(t_i) / S(t_{i-1})), t_i)$ in a window. Quantity $S(t_i)$ is the price of security at time t_i and $S(t_0)$ is the initial price of the security at t_0 .

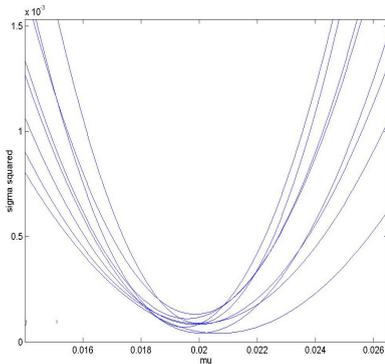


Figure 7: Plot of simulated $f(\mu)$ against μ , with assumed mean, $\mu = 0.02$, $\sigma = 0.05$.

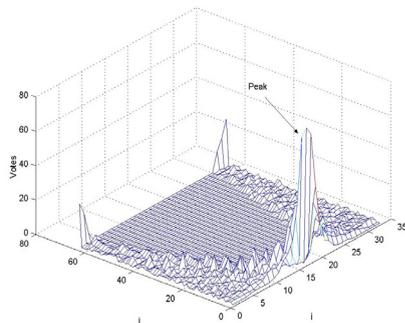


Figure 8: A 3-dimensional accumulator array for sigma-quadratic equation.

We consider short-term sample stock price logarithms $[\log(S(t_0)), \log(S(t_1)), \dots, \log(S(t_N))]$ where $\log(S(t_i))$ is assumed at times $t_i \in [0, 1, 2, \dots, N]$. Taking

window-data (window-size n , $0 < n < N$) from the sample and of duration N , we get $f_1(\mu, \sigma^2)$, $f_2(\mu, \sigma^2)$, \dots , $f_k(\mu, \sigma^2)$ (k windows of n points, so that $nk = N$) from each window. We simulate 50 price histories with window size, $n = 5$, and produce a family of graphs as depicted in Figure 7. Figure 8 depicts the accumulator array of the adapted Hough transform. The point with the brightest colour in the accumulator indicates the point of concentration of the curves, i.e. the point where there are many voting curves. In Fig. 8 the highest point of concentration has white colour as shown in the colour bar.

The curves tend to meet at a point close to the assumed mean. A 3-dimensional mesh accumulator array of Figure 7 is shown in Figure 8 with the highest ordinate representing the maximum votes in the cell where the curves meet. The mesh accumulator array is shown as a butterfly-type shape with a large peak and using-like ridges around it (Leavers [9]). The heights of the ordinates are proportional to the number of votes in each cell.

3.2 Grey scale testing on geometric Brownian motion

In this simulation we use a simulated GBM history with $N = 400$, $\mu = 0.020$ and $\sigma^2 = 0.0025$. Figure 9 depicts the logarithms of asset prices with 95% ($1.96\sigma\sqrt{t}$) envelope about the trend line $\log P = 2 + \mu t$. Figure 10 depicts the 3-dimensional accumulator array for the voting curves with the cell that contains most votes shown by the highest peak.

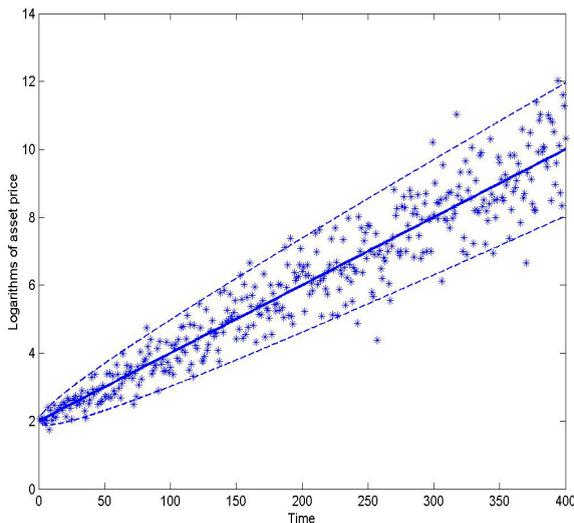


Figure 9: Logarithms of asset prices versus time, t , with simulated mean, 0.020 and volatility, $\sigma^2 = 0.0025$ and $\pm 1.96\sigma\sqrt{t}$ envelope (dotted lines) about the trend line (bold line).



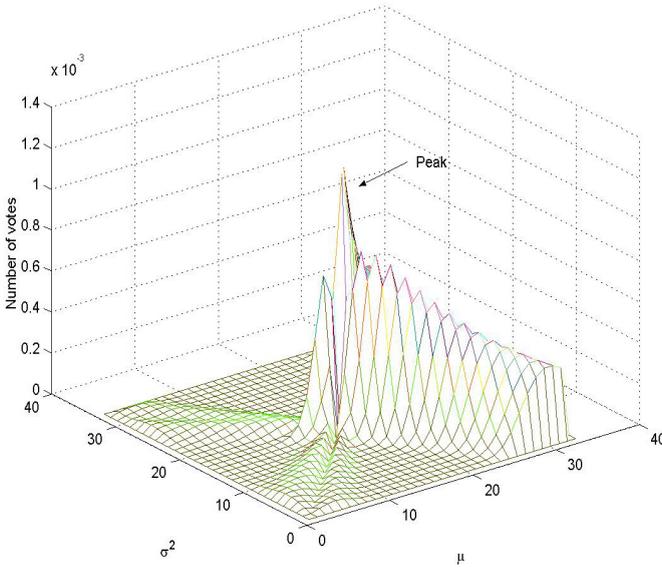


Figure 10: A 3-dimensional grey-scale accumulator array with $N = 400$, $\mu = 0.020$ and $\sigma^2 = 0.0025$ with labelled peak corresponding to a peak.

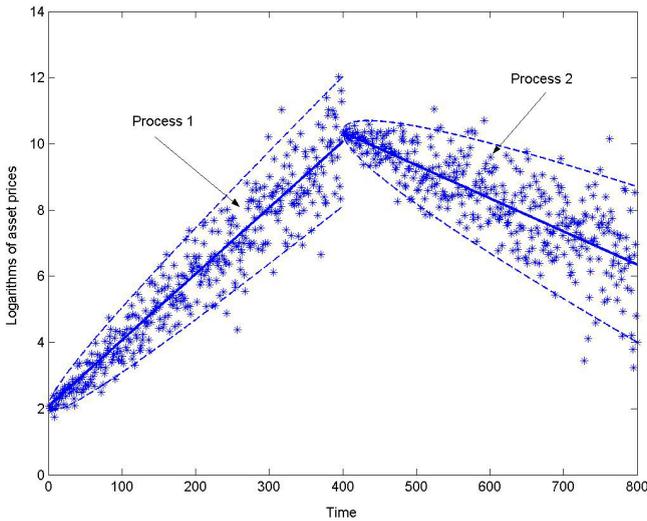


Figure 11: Logarithms of simulated asset prices where two processes are shown with $\pm 1.96\sigma\sqrt{t}$ envelope (dotted curves) about the trend lines (bold lines), with a sudden change of GBM process at time 401.



3.3 Simulation of multiple processes in geometric Brownian motion

We simulated asset prices with $N_1 = 400$, $N_2 = 400$, $\mu_1 = 0.020$, $\mu_2 = -0.010$ and $\sigma_1 = 0.05$, $\sigma_2 = 0.06$. Figure 12 depicts the logarithms of the asset prices plotted against time, showing the two processes. Figure 13 depicts the grey-scale accumulator array where two peaks corresponding to the two processes are clearly shown.

In this section, we have tested for geometric Brownian motion and found out that processes can be detected using grey scale voting techniques. Peaks in the three-dimensional accumulator array represent processes in the price history space. This technique is robust to both impulsive noise and pixel occlusions. In the above cases we have discussed with linear geometric Brownian motion. In the next section we apply the HT to real-world market price data.

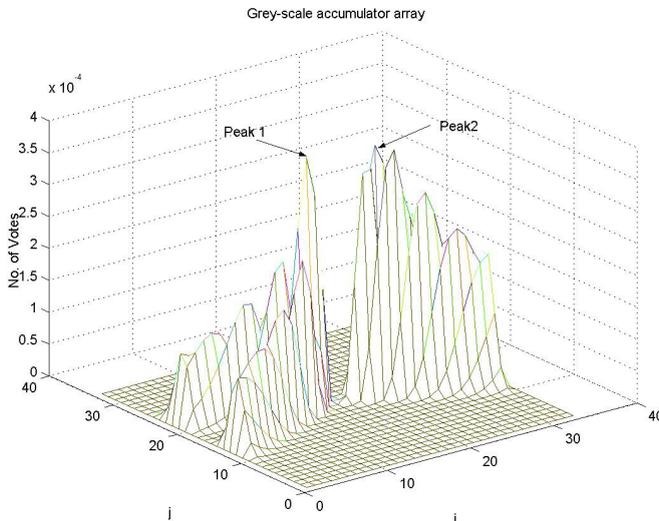


Figure 12: A 3-dimensional grey-scale accumulator array showing two peaks. The peaks labelled Peak 1 and Peak 2 correspond to the two processes in Figure 11.

4 Pattern on real world market data

In this section we use two examples to test the Hough transforms on geometric Brownian motion with real world market data.

4.1 Example 1: Manchester United Football Club (MANU) (upward trend)

We consider a 64-point price history of Manchester United Football Club (MANU) from 12 Jul.2003 to 13 Oct. 2003 in which there is a positive market



trend. Fig. 13 depicts a 95% envelope about the trend line where most of the price-history points are included within the envelope. Figure 14 depicts 3-dimensional grey-scale accumulator array with a single peak well delineated.

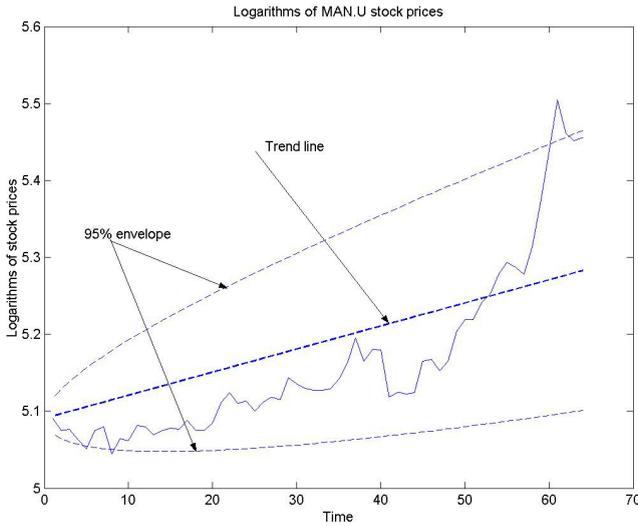


Figure 13: Share price curve with 95% envelope from estimate GBM parameters $\hat{\mu} = 0.0030, \hat{\sigma} = 0.0116$ superimposed.

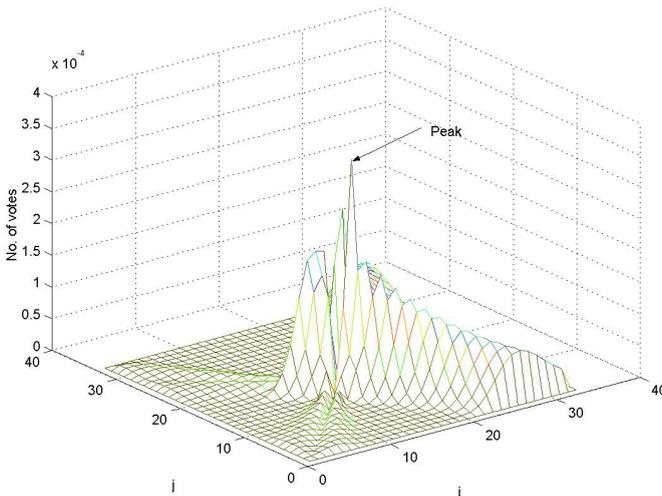


Figure 14: 3-dimensional grey-scale accumulator array for volatility curves of Manchester United Football Club (MANU) with one peak clearly shown.

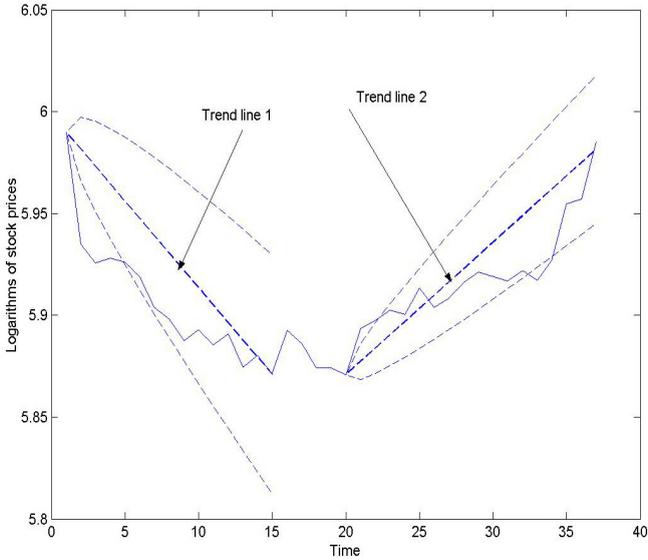


Figure 15: Logarithm curve of share prices of Scottish and Newcastle Brewers Plc, with 95% envelopes (dotted curves) about the trend lines and $\mu_1 = -0.0085$, $\mu_2 = 0.065$, $\sigma_1 = 0.0080$, $\sigma_2 = 0.0045$.

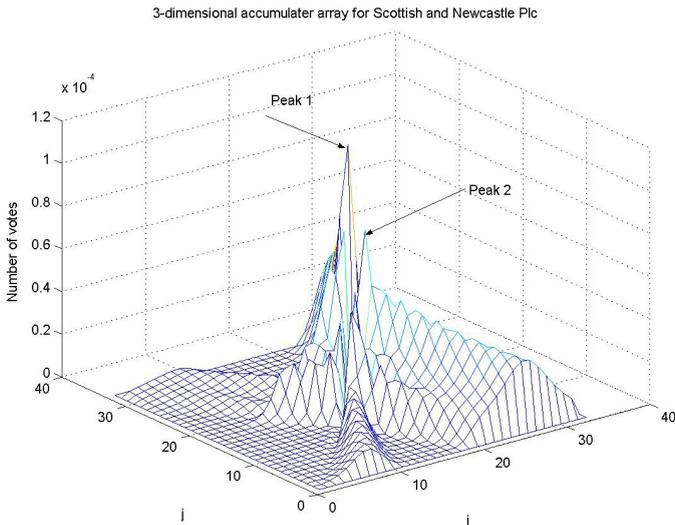


Figure 16: 3-dimensional grey-scale accumulator array for Scottish and Newcastle Brewers Plc. From the figure two peaks, peak 1 and peak2, are clearly shown.

4.2 Example 2: Scottish and Newcastle Brewers Plc (sudden change of trend)

We use a 37-point price history of brewers Scottish and Newcastle Brewers Plc during the period 20th Jun. To 8th Aug. 2003 in which there is a sudden change of market trend. We show the presence of two processes in the market. High peaks, Peak 1 and Peak 2 in the 3-dimensional grey-scale accumulator array represent these two processes. Figure 15 depicts asset price of the thirty-seven real asset prices. Figure 16 depicts the corresponding 3-dimensional grey-scale accumulator array with peaks, Peak 1 and Peak 2, clearly delineated.

5 Conclusion

In this paper, we have shown that AI can be adapted to detect stochastic process in simulated and real market data involving geometric Brownian motion. This technique has been shown to be robust to market data occlusion, noise in data and is able to detect multiple processes in a market. An additional virtue of this technique is its ability to communicate complex ideas with tremendous simplicity, clarity, speed and power. It also supports effective or rapid decision making in the market. As is evident in the market, velocity of change is increasing, amidst greater complexity and chaos. So, possessing a deep understanding of the market patterns can enable critical decision making whenever trading is to be undertaken an urgent issue-this helps the trader to decide when to trade and when not to trade. For financial management, knowledge of market patterns will also help in the quality of decision to be made.

References

- [1] Aguado, A. S., E. Montiel and M. S. Nixon (2000). Error Analysis of the Generalised Hough Transform. *Journal of Mathematical Imaging and Vision* 12(1), pp. 25 – 42.
- [2] Bachelier, L. (1900). *Théorie de la Spéculation*. Annales Scientifique de L'École Norm. Sup. 11-17, pp. 21-28.
- [3] Black, F. and M. Scholes (1973). *The pricing of options and Corporate Liabilities*. Journal of Political Economy 18, pp. 637 – 659.
- [4] Boyle, P.P. and Thangaraj (2000). *Volatility estimation from observed option prices*. Decisions in Economics and Finance 23 (1), pp. 31 – 52.
- [5] Duda, R. O. and P. E. Hart (1972). The use of Hough Transform to detect lines and curves in pictures. *Comm-ACM* 15, pp. 11 – 15.
- [6] Flint, A. D., M. Ingleby and D, A, Morton (1992). A new generalization of the Hough transform – applied to automatic prognosis of condition from monitored data. Proceedings, IEEE Conference intelligence 92, pp. 261 –267.
- [7] Hough, P. V. (1962). *Method and means for identifying complex patterns*. U.S. Patent No. 3069654.



- [8] Kingdon, J. (1997). *Intelligent Systems and Financial Forecasting*. Springer-Verlag: London.
- [9] Leavers, V. P.(1992) *Shape detection in Computer Vision using Hough Transform*, Springer-Verlag, Berlin.
- [10] McDonald, J. (2001) *Detecting and tracking road markings using Hough transform*. Proceeding of the Irish Machine Vision and Image Processing Conference, 2001. National University of Ireland, Ireland
- [11] Merton, R. (1973) *The theory of Rational Option Pricing*, The Bell Journal of Economics and Management Science 4,pp. 141-183.
- [12] Onyango, S. N. (2003). *Extracting Stochastic Process Parameters from market price data: A pattern recognition approach*. Unpublished PhD Thesis, University of Huddersfield, Huddersfield, UK.
- [13] Princen, J. J., J. Illingworth and J. Kitler (1992). *A formal definition of the Hough transform: properties and relationships*. Journal of Mathematics Imaging and Vision 1, pp. 153 –168.
- [14] Toft, P. (1996). *Radon Transform: Theory and Implementation*. Unpublished PhD Thesis. Department of Mathematical Modelling, Technical University of Denmark.
- [15] Weik, M. (1989). *Communications Standard Dictionary*, 2nd Ed. Van Nostrand Reinhold Company: New York.
- [16] Zaprakis, A. and A. Refenes (1999). *Principles of Neural Model Identification, Selection and Adequacy with Applications to Financial Econometrics*. Springer-Verlag: London.

