A General Stability Analysis on Regional and National Voting Schemes Against Noise

—Why is an electoral college more stable than a direct popular election?

Liang Chen

Computer Science Program, University of Northern British Columbia,

Prince George, B.C., Canada V2N 4Z9

E-mail: lchen@ieee.org

Naoyuki Tokuda

SunFlare Company, Shinjuku Hirose Bldg, Yotsuya 4-7, Shinjuku-ku,

Tokyo, Japan 160-0004

Email: tokuda_n@sunflare.co.jp

Abstract

By relaxing the restrictive weak average distribution assumption, we have developed a newgeneral stability analysis on the national and the regional voting which remains valid over the entire range in size of the partitioned regions. Our analysis demonstrates that the regional voting (electoral college, for example) is always more stable than the national voting (known as “direct popular voting” in political science), and that the stability margin of the regional voting always increases as the size of such partitioned regions decreases up to a certain limit beyond which the stability margin starts to decrease, asymptoting to a national voting limit where the unit of voting cell approaches the size of the partitioned region so that the improved stability of the regional voting by localizing the effects of noise into a restricted number of smaller effective area will be not be effective.. We show that the regional voting asymptotes to the national voting in two extreme limiting cases when the region size not only decreases to a voting cell size but also increases to an original national voting size.

Index Terms

national voting, regional voting, image processing, stability

I. Introduction

Among many decision making schemes starting with more casual paper-rock-scissor, coin tossing, , and more serious SUN Tzu’s art of war [1], voting used in candidate selection such as the US presidential election constitutes one of the most important decision making processes we resort to not only in our daily social life but also in many scientific studies.

In daily life, the so-called national and the regional voting are the two most popular voting schemes. We resort to a national voting, also known as a direct popular election, in selecting a mayor of our local communities, for example, where we select the winner from among candidates directly by a simple majority of the entire voting population of the nation. The regional voting on the other hand, involves a more complex process whereby winners are determined by the majority of the winning regions into which the entire national voting district is divided, as into prefectures or states; the US presidential election, also being called as electoral college, is a most well known example of regional voting scheme. The 2000 US presidential election has rekindled interest in exploring the election system selection in political science. Is there any advantage for the US presidential election to use electoral college rather than a simpler, more direct popular election? It is worthwhile to review how the voting schemes have been used in pattern recognition and machine vision here. Consider the case of national voting scheme : A pixel-based matching approach could be employed to define the similarities between frames in a video sequence so as to construct video loops with apparently natural variation; A perceptron that can be used for the recognition of translation initiation sites in DNA sequences is another example where a weighted national voting approach is taken. The regional voting scheme has also many applications in scientific research. For example, in content-based image retrieval, the image matching is always done by using more selective features from partitioned images rather than informative features extracted from the entire image without subdivision, i.e., "national" type features [2]. We see that different regions give rise to different hypotheses, so that by combining all hypotheses rooted in all regions we shall then come up with an overall decision. One of the simple partitioning is to divide the image into blocks of equal size and "summarize" the dominant feature values, which could be taken as "regional" features, in each block to be used for matching purpose [3]. Keren gives an interesting example of image " style" classification by regional voting approach [4] where he shows how to determine the author of an image; by firstly extracting, from each possible 9×9 sized block in the image, the features in the form of DCT coefficients, then labeling the block by comparing the Bayes possibilities of the block belonging to different artists, and finally determining the style of the entire image by a majority voting[1]. Experiments [4] showed that Keren's method is able to yield results which agree very well with human intuition. Keren also noticed that classification performance deteriorates significantly when the size of the blocks is increased to 18×18 [2] . While many approaches in pattern recognition and machine vision can be viewed as regional or national voting procedures, it makes sense to see which voting scheme performs better for pattern recognition. We believe that there is a keen need for a proper explanation on the advantages / disadvantages of using regional / national voting style approaches, and on how the size of a "block" in regional voting affects the voting/matching performances.

While voting theorists in social science field are more concerned with the "fairness" of social

[1][4] uses a sliding window of size 9×9 to get the blocks so that the number of blocks used in voting is very close to the number of pixels in the image, implying that Karen's voting scheme is almost equivalent to the extended regional voting which we will discuss in Section III-C

[2]Keren guesses that 5x5 and 3x3 would be too small to be useful for his image "style" classification, although he has not done any experiment on these sizes (private communication).

vote systems [5], [6] , such as universality criterion, monotonicity criterion, independence of irrelevant alternatives criterion, citizen sovereignty criterion, and non-dictatorship criterion, to name a few [5], [6] [3]. But it is known that one criterion always violates another one or more, and Arrow's impossibility theory [5] shows that there is no 'perfect' voting scheme that meets all these criteria, implying that the selection of a voting scheme is, by necessity, subjective.

We could see that, the fairness of the above criteria comes from the fact that they could be used to reflect the democracy of a society that each voter and each candidate share and retain the same amount of rights. But the criteria needed in democratic election systems are not much of our concern here in scientific research such as pattern recognition and image retrieval, because our paper is concerened with the stability against noise so that the concept of fairness is not directly relevant to the cells or pixels.Very little has been done on the stability analysis of decision-making process in political science. Only some experimental works are available showing that some special voting scheme is the more reliable or more stable than the others depending on specific applications [7], [8], [9].

We have made some progress on the stability analysis on the national and the regional voting [10], [11]. In [10], we succeeded in analyzing the stability of the voting process subject to an Average Distribution Assumption where we set up a discrete model on a simplest possible voting scheme with only two nearly competitive candidates with a nearly equal number of supporters. We have shown that the regional voting is always more stable than the national voting against the concentrated noise which is applicable to image processing applications and that the stability of the regional voting should increase as the size of subdivided regions decreases, as long as the size of regions is *not* too small to accommodate the Average Distribution Assumption [10]. The concentrated noise treated in the image processing applies to damaged or polluted images due to ink blotting, leaking sunlight through sunshade, carriers, or due to the electric noise at a certain period of times. The Average Distribution Assumption assumes homogeneity in

[3]These criteria guarantee, respectively, that voting should always find a winner between two alternatives or say that there is a social indifference between them, that societal preference order should change only by (possibly) promoting the same option if one individual voter modifies her or his preference order by promoting a certain option, that societal preference of two alternatives should be independent of preferences for other options, that every possible societal preference order should be able to be achieved by some set of individual preference ballots, and that societal preferences should not depend only upon the preferences of one individual "dictator".

distribution of supporters where the distribution of each region follows that of the whole nation. By extending the analysis to a continuous model where we allow the presence of both white noise and concentrated noise [11], we have succeeded in *relaxing* the average distribution assumption to a Weak Average Distribution Assumption which assumes the dominating relationship between supporters of candidates in each of the partitioned regions and in the nation, and showed that the conclusion on the improved stability of regional voting over national voting and that of the regional voting with smaller sized regions over the regional voting with larger sized regions remains valid as long as the Weak Average Distribution Assumption is valid.

The average distribution assumption and the weak average distribution assumption share the relationship of consistency[4]. We believe the (weak) consistency assumption should hold for many cases, but the validity may not extend to all the situations, especially to regions involving very few voters, say only 2 or 3 voters, or even only 1 voter.On the other hand, our experiments on facial recognition in both [10] and [11] convincingly demonstrate that the "stability" of the regional voting decreases as the size of the regions decreases in a "not too small" range in region size. How can we set up an analytical model for the situation? This paper is motivated to provide a general model capable of explaining the stability behavior that has not been explained by previous models by relaxing the restricted concept of (weak) average distribution assumption.

To relate the present analysis to practical applications in pattern matching , we need to discuss the registration of objects for recognition. Pattern matching or recognition by pixel-based (or single-voter based) vote counting encounters many difficulties if the registration of an unknown object for recognition and that in the training set is imperfect, particularly . so when the differences are large. Under the situation, it is not useful to compare the features or color values of an isolated pixel / cell on a pixel-to-pixel / cell-to-cell basis for object recognition. One effective well known scheme of proper registration is the PCA approach by an SVD decomposition [13]; and it has been very successfully used for facial matching problem.This explains why the features extraction schemes such as principle component analysis (PCA), Fourier transformations, Gabor transforms and et. al play a more important role. These schemes are developed to extract the characteristics features from the sets of pixels. The feature extraction process always exploits

[4]The Weak Average Distribution Assumption is one of the most important, major criteria in evaluating a voting system in political science, where it is called `Consistency`. If the voting system is not consistent, it may be easily "manipulated" by government by constructing the strategically configured election districts [12].

the well-known dimension-reducing schemes. Let G represent such a "global matching" process as PCA schemes or Fourier Transforms which involves a more general process than the vote counting, allowing the vote counting at its simplest case . In the regional matching scheme $r(G)$, we make a final decision by a simple majority of the winning regions using the "winner-take-all" principle within all the regions of the pre-partitioned nation. Here the winner of each pre-partitioned region is determined by employing G on the region. Although our model, and the conclusion thereof on stability of national and regional voting schemes obtained thereafter, are based on the analysis on the simplest vote counting scheme, we expect the model remains valid for the general national matching schemes G and their regional matching versions $r(G)$. Consequently, we will pose a conjecture in this paper that our model is valid in general decision-making processes, where feature extraction schemes are used in place of simple vote counting for determining a winner in national matching and also in each region for regional matching. An experimental verification of our theory presented in Section IV-C.2 is carried out by exploiting the PCA approach in both national matching scheme G, and its regional matching version $r(G)$.

The rest of the paper is constructed as follows: We introduce the basic terms used in Section II. The main theorems regarding the stability are given in Section III. We list the conclusions derived from the main theorems on improved performance of regional voting and its relationship to region sizes are given in Section IV, along with the conjecture on the validity of our model for more general matching processes using complex feature matching schemes in place of simple voting counting for determining the winners in nation for national matching and also in each region for regional matching. The experimental verifications of our conclusions are also given in Section IV. The discussions on the applicability of the model and future work is given in Section V.

II. Basic Model & Assumption

A. Basic Model

1) We simplify the voting model to a simplest two candidate system of A and B. The nation is represented by a rectangle area comprising $l \times w = N$ (l and w being positive integers) unit cells, where a cell is the basic unit of votes, that is, one cell comprises one vote. The voting for candidates *A* and *B* of a cell, is defined as a selection either for *A* or *B*. Let α and β denote the percentage of the (total) votes which candidates *A* and *B* get from

the whole nation in the absence of noise; we have $\alpha + \beta = 1$; In the absence of noise candidate A wins in the national voting, so that we have $1 > \alpha > \beta > 0$.[5]

As we consider the voting in a nation having a reasonable size of voting population, we could regard α (or β) as the probability of the votes for candidate A (or B) when we arbitrarily choose one voter in the nation. This implies that in a large subset of the entire nation, the rate of A supporters should maintain the same rate throughout the nation [6].

2) For regional voting purpose, the rectangular nation is divided into equal shaped rectangles, called regions of size $r_l \times r_w$ (r_l and r_w being positive integers), where l and w are divisible by r_l and r_w independently.
3) The national voting is implemented over the entire nation, a winner in the national voting is decided by a simple majority of the votes throughout the nation; a winner in the regional voting is decided by the "winner-take-all" principle namely by a majority of the winning regions, where the winner of each region is determined by a simple majority of votes within the region.
4) The noise is defined as a change of environment that enforces a change of voting result. When subjected to noise, the votes of some of the cells will undergo a change from A to B, some from B to A, and others may remain unchanged. The noise that influence votes to change from A to B (or B to A) is called anti-A-noise (or anti-B-noise). A vote which undergoes a change from A to B (B to A) is called an anti-A-noise-contaminated vote (anti-B-noise-contaminated vote).
5) Two types of noise caused by independent, known or unknown, sources are considered; concentrated noise which influences the votes within a concentrated block(s) of cells, and white noise that is distributed uniformly over the whole nation randomly.
6) A set of anti-A-white noise (or anti-B-white noise) is dispersed uniformly over the nation, producing a uniform chance of converting A to B (or B to A). The result of white noise thus could be regarded as a change in the probability of voting for A from α to a new value, and a change in probability of voting for B from β to a new value accordingly[7].

[5]In the later section, we will see that A will also win the regional voting in a noise free environment

[6]Note that we do *not* regard a region as a large or "large enough" subset of the nation any longer in the analysis so that a weak average distribution assumption is no longer needed.

[7]It is obvious that the union of a set of white noise is also a set of white noise.

7) A set of anti-A-concentrated noise (or anti-B-concentrated noise) is defined as the union of non-overlapped rectangle blocks of size $n_l \times n_w$, on each of which all the votes for A (or B) will be changed to B (or A). The corresponding union of these rectangle areas is called a noise concentrated area, and $n_l \times n_w$ is called the size of noise blocks[8].
8) In accordance with the above two types of noise, the anti-A-noise-contaminated votes (anti-B-noise-contaminated votes) comprise also two different types depending on the noise type, namely anti-A-concentrated-noise-contaminated votes (anti-B-concentrated-noise-contaminated votes) and the anti-A-white-noise-contaminated votes (anti-B-white-noise-contaminated votes).

   Notice that, when both of white noise and concentrated noise coexist, some noise-contaminated votes may belong to both of these two types, as it will be seen in the proof of theorem 1. As white noise is dispersed uniformly over the nation, the ratio of A (and that of B) supporters subjected to white noise in an arbitrarily chosen large area of the nation should be equal to that in the whole nation. This implies that a percentage of white noise contaminated votes in the noise concentrated area should not change from that in the whole nation. This is useful in computing an overlap of white noise contaminated votes and concentrated noise contaminated votes. It is understandable as the noise concentrated area can be considered as reasonably large.
9) We call a region concentrated noise polluted *if and only if* the conjunction set of the region and the noise-concentrated area is not empty.
10) Because we are interested in computing the lower bounds of the voting stability throughout this paper, we consider only the anti-A noise in the analysis. Thus anti-A noise, anti-A-concentrated noise, anti-A-white noise, anti-A noise contaminated votes are referred to as noise, concentrated noise, white noise, or contaminated votes hereinafter, respectively.
11) $\aleph_c$, and $\aleph_w$ denote the number of concentrated-noise contaminated votes, and the number of white-noise contaminated votes separately.

The two-candidate election system we adopt in this paper is not as restrictive as it looks, because as we show in Section V-A.2 the result of the conclusion is still valid for a multi-candidate model

[8] Intuitively, the "white noise" is isolated and scattered randomly over discrete "points" of the nation while "concentrated noise" is distributed over connected, continuous areas which may be randomly distributed across the nation.

involving three or more candidates.

B. Assumption

We always assume that both the voting population of a nation and the amount of noise are large so that both the total number of noise contaminated votes and the noise concentrated area are large. It is important to note that we no longer assume the voting population of subdivided regions to be large compared with the nationś population.

Assumption 1 (Basic Assumption): In the absence of concentrated noise, the percentage of A (or B) dominated regions among all the regions, or among a set of a large number of arbitrarily chosen regions which may or may not be neighbored, is equivalent to the chance that A gets more votes than B does (or B gets more votes than A does) in any region.

The word "arbitrarily" in choosing regions here, and also in the lemmas below, is used to indicate that a selection process / approach is independent of any knowledge on the specific distribution of A or B supports on the chosen regions other than that of the whole nation, and also independent of the sources of white noise if white noise is present.

This assumption implies that:

Lemma 1: In the absence of both white and concentrated noise, among a set of a large number of arbitrarily chosen regions in the nation, the proportion of A and B dominated regions P_A and P_B can be computed by:

$$P_A = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \binom{r_l r_w}{y} \beta^y \alpha^{r_l r_w - y}$$

$$P_B = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \binom{r_l r_w}{y} \alpha^y \beta^{r_l r_w - y}$$

Together with the definition of white noise in previous section, the Basic Assumption also implies that:

Lemma 2: In the presence of only white noise, in a set of a large number of arbitrarily chosen regions in the nation, the proportion of A and B dominated regions, P'_A and P'_B, can be computed by:

$$P'_A = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \binom{r_l r_w}{y} \beta'^y \alpha'^{r_l r_w - y}$$

$$P'_{_B} = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \begin{pmatrix} r_l r_w \\ y \end{pmatrix} \alpha'^{y} \beta'^{r_l r_w - y}$$

where, α' and β' are the percentage of the votes for A and for B, in the presence of white noise.

A region is referred to pro A (A dominated) or pro B (B dominated) if A dominates B in the region or B dominates A.

Of course, we have $P_{_A} + P_{_B} = 1$, and $P'_{_A} + P'_{_B} = 1$.

We could notice that $\frac{P_A}{P_B} \to \infty$, if $r_l r_w \to \infty$ and $\alpha > \beta$; this is what the Weak Average Distribution Assumption says. But here we *cannot* claim that $r_l r_w \to \infty$ is always true, as the size of the regions can be *small*.

The fact that $P_{_A} > P_{_B}$ iff $\alpha > \beta$, simply says in noise free environment, both of national and regional voting select the same candidate.

We should regard the set of concentrated noise polluted regions as a set of "arbitrarily" chosen regions discussed in the above Basic Assumption, because concentrated noise also have to be "random", although they occur in the form of concentrated blocks. This is to say, the above lemmas should be valid for the set of concentrated noise polluted regions. Thus, although among the concentrated noise polluted regions, some of them should be originally "A" dominating and some originally "B" dominating, the proportion of originally "A" dominating regions in all the concentrated noise polluted regions equals that of all the regions of the nation.

III. Main Theorems

A. The Stability of National Voting

Theorem 1: The national voting will preserve the original candidate A iff

$$\aleph_c + \aleph_w \times \frac{N - \aleph_c/\alpha}{N} < \frac{\alpha - \beta}{2} \times N.$$

Proof: Given the number of anti-A-concentrated-noise-contaminated votes $\aleph_c$, the size of anti-A-noise-concentrated area can easily be calculated as $\aleph_c/\alpha$. Among $\aleph_w$ anti-A-white-noise-contaminated votes, $\frac{\aleph_c/\alpha}{N} \times \aleph_w$ votes come from the anti-A-Noise-Concentrated area. It is easy to see that $\frac{\aleph_c/\alpha}{N} \times \aleph_w$ votes are overlapped between $\aleph_c$ anti-A-concentrated-noise-contaminated votes and $\aleph_w$ anti-A-white-noise-contaminated votes. The national voting is able to preserve the original candidate selection, if and only if the number of overall anti-A-concentrated-noise-contaminated votes is less then $\frac{\alpha-\beta}{2} \times N$. ■

B. The Stability of Regional Voting

Theorem 2:

1.The original candidate selection of the regional voting will be retained if:

$$\aleph_c < \frac{\frac{n_l n_w}{r_l r_w}}{(\left\lceil \frac{n_l-1}{r_l} \right\rceil + 1)(\left\lceil \frac{n_w-1}{r_w} \right\rceil + 1)} \cdot \frac{P_A(\aleph_w) - P_B(\aleph_w)}{1 + P_A(\aleph_w) - P_B(\aleph_w)} \cdot \alpha \cdot N,$$

and

$$\aleph_w < (\alpha - \beta)/2 \times N.$$

2. The candidate selection of regional voting could still be preserved on lucky cases when

$$\frac{\frac{n_l n_w}{r_l r_w}}{(\left\lceil \frac{n_l-1}{r_l} \right\rceil + 1)(\left\lceil \frac{n_w-1}{r_w} \right\rceil + 1)} \cdot \frac{P_A(\aleph_w) - P_B(\aleph_w)}{1 + P_A(\aleph_w) - P_B(\aleph_w)} \cdot \alpha \cdot N \leq \aleph_c < \frac{P_A(\aleph_w) - P_B(\aleph_w)}{1 + P_A(\aleph_w) - P_B(\aleph_w)} \cdot \alpha \cdot N;$$

and

$$\aleph_w < (\alpha - \beta)/2 \times N,$$

where $P_A(\aleph_w)$ and $P_B(\aleph_w)$ denote the percentages of A-pro and B-pro regions under white noise, and can be calculated as

$$P_A(\aleph_w) = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \begin{pmatrix} r_l r_w \\ y \end{pmatrix} (\beta + \aleph_w/N)^y (\alpha - \aleph_w/N)^{r_l r_w - y},$$

$$P_B(\aleph_w) = \sum_{y=0}^{\lfloor \frac{r_l r_w}{2} \rfloor} \begin{pmatrix} r_l r_w \\ y \end{pmatrix} (\alpha - \aleph_w/N)^y (\beta + \aleph_w/N)^{r_l r_w - y}.$$

Proof: If $\aleph_w < (\alpha - \beta)/2 \times N$, it is easy to see that $P_A(\aleph_w) > P_B(\aleph_w)$, and thus the regional voting is able to preserve the original candidate selection In the absence of concentrated noise.

1. Suppose that the rate of concentrated noise polluted regions is X $(0 \leq X \leq 1)$. According to Lemma 2, we could know that $P_A(\aleph_w)$ portion of these concentrated noise polluted regions are Pro-A regions and $P_B(\aleph_w)$ portion are Pro-B regions originally in the presence of white noise only. Then when $X + (1-X)P_B(\aleph_w) < (1-X)P_A(\aleph_w)$, i.e., $X < \frac{P_A(\aleph_w) - P_B(\aleph_w)}{1 + P_A(\aleph_w) - P_B(\aleph_w)}$, the number of Pro-A regions should be still larger than the number of Pro-B regions, even if we

regard all concentrated noise polluted regions as Pro-B regions, while actually in all possibility some of them still remain Pro-A in practice. Thus, the regional voting should be able to preserve the original voting selection when $X < \frac{P_A(\aleph_w)-P_B(\aleph_w)}{1+P_A(\aleph_w)-P_B(\aleph_w)}$.

Noticing that a noise block of size $n_l \times n_w$ can be partitioned into at most $(\lceil \frac{n_l-1}{r_l} \rceil + 1)(\lceil \frac{n_w-1}{r_w} \rceil + 1)$ different regions, we have:

$$\frac{S_r}{S_c} \le (\left\lceil \frac{n_l - 1}{r_l} \right\rceil + 1)(\left\lceil \frac{n_w - 1}{r_w} \right\rceil + 1)\frac{r_l r_w}{n_l n_w}.$$

where $S_r = X \cdot N$ is the total size of concentrated noise polluted regions within the nation, and S_c denotes the total size of noise concentrated area. To meet the requirement on X of $X < \frac{P_A(\aleph_w)-P_B(\aleph_w)}{1+P_A(\aleph_w)-P_B(\aleph_w)}$, we only need the following inequality:

$$|S_c| < \frac{\frac{n_l n_w}{r_l r_w}}{(\left\lceil \frac{n_l-1}{r_l} \right\rceil + 1)(\left\lceil \frac{n_w-1}{r_w} \right\rceil + 1)} \cdot \frac{P_A(\aleph_w) - P_B(\aleph_w)}{1 + P_A(\aleph_w) - P_B(\aleph_w)} \cdot N.$$

Substituting the relation of $\aleph_c = |S_c| \times \alpha$, the conclusion of item 1 of Theorem 2 follows.

2. To prove item 2 of Theorem 2, it is sufficient to consider the case when the size of the noise concentrated area equals the total size of concentrated noise polluted regions. ■

C. *Extended Regional Voting*

In Theorem 2 above, the ceiling operations are used to develop a sufficient condition of stability that constitutes a worst possible condition whereby each of the noise blocks contaminates a maximum number of regions. It is unlikely that the worst situation for each noise block happens at the same time. Some appropriate averaging will be introduced here by shifting the partitions with noise distribution fixed.

Geometrically, we could regard the pair of the opposing edges along the outer boundary of the rectangular nation glued so that they are to glide onto the other end as we move across the boundary allowing a total of $r_l \times r_w$ different partitions by merely shifting all of the horizontal and vertical boundaries of the regions. Taking into consideration of all the $r_l r_w$ different partitions which result as a total of $r_l r_w \times N/(r_l r_w) = N$ different regions, *Extended Regional Voting* is defined as the regional voting where the winner is selected by a majority of the winning regions from among the N regions rather than from $K = N/(r_l r_w)$ regions as a result of shifting operations. The following theorem shows the improved result:

Theorem 3: The extended regional voting will retain the candidate selection if:

$$\aleph_c < \frac{n_l n_w}{(r_l + n_l - 1)(r_w + n_w - 1)} \cdot \frac{P_A(\aleph_w) - P_B(\aleph_w)}{1 + P_A(\aleph_w) - P_B(\aleph_w)} \cdot \alpha \cdot N$$

and

$$\aleph_w < \frac{1}{2}(\alpha - \beta)N$$

where $P_A(\aleph_w)$ and $P_B(\aleph_w)$ are calculated as they are in Theorem 2.

Proof: To prove the theorem, we must show that, among all the possible $r_l r_w$ different partitions, each of the $n_l \times n_w$ sized noise block is capable of contaminating a total of $(n_l + r_l - 1)(n_w + r_w - 1)$ different regions. We only consider the case of $n_l \leq r_l$ and $n_w \geq_w$ here. For an $n_l \times n_w$ sized noise block, among all the possible $r_l \times r_w$ partitions, there are $r_l - n_l + 1$ partitions, that divide the block into $\left\lceil \frac{n_w}{r_w} \right\rceil$ different regions, and also $r_l - n_l + 1$ partitions that divide the block into $1 + \left\lceil \frac{n_w - 1}{r_w} \right\rceil$ different regions, $\cdots$, $r_l - n_l + 1$ partitions divides it into $\cdots$, $1 + \left\lceil \frac{n_w - r_w + 1}{r_w} \right\rceil$ different regions. We can find another $n_l - 1$ partitions where each can divide the block into $2\left\lceil \frac{n_w}{r_w} \right\rceil$ different regions, and $n_l - 1$ partitions each can divide the block into $2 + 2\left\lceil \frac{n_w - 1}{r_w} \right\rceil$ different regions, $\cdots$, $n_l - 1$ partitions divide the block into $2 + 2\left\lceil \frac{n_w - r_w + 1}{r_w} \right\rceil$ different regions. Summing up all the possible terms above, a noise block of size $n_l \times n_w$ will be divided into $(n_l + r_l - 1)(n_w + r_w - 1)$ different regions for all the $r_l \times r_w$ different partitions. Fig. 1 illustrates the case of $n_l = 5 < r_l = 9$ and $n_w = 11 > r_w = 4$. Shifting the partitions from bottom up, we see immediately that the partitions through the solid dots of the first line divide the noise block into $\left\lceil \frac{n_w}{r_w} \right\rceil$ regions, while those of the second line divide the noise block into $1 + \left\lceil \frac{n_w - 1}{r_w} \right\rceil$ regions, $\cdots$; the partitions through the crossed points of the first line divide the noise block into $2\left\lceil \frac{n_w}{r_w} \right\rceil$ regions, while those of the second line divide the noise block into $2 + 2\left\lceil \frac{n_w - 1}{r_w} \right\rceil$ regions, $\cdots$.

Let the total size of noise concentrated area be S_c. Now, counting all the polluted regions by the extended regional voting, the total size of all the concentrated noise polluted regions S_r can be computed by:

$$\frac{|S_r|}{|S_c|} \leq \frac{(n_l + r_l - 1)(n_w + r_w - 1) \times r_l r_w}{n_l n_w},$$

Let X ($0 \leq X \leq 1$) be the rate of concentrated noise polluted regions among all the N regions of this extended regional voting. As shown in the proof for Theorem 2, when $X + (1-X)P_B(\aleph_w) < (1 - X)P_A(\aleph_w)$, i.e., $X < \frac{P_A(\aleph_w) - P_B(\aleph_w)}{1 + P_A(\aleph_w) - P_B(\aleph_w)}$, the number of Pro-A regions should be still larger

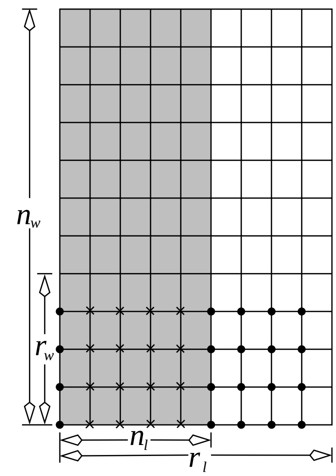


Fig. 1. Different Partitions in the Nation

than the number of Pro-B regions, even if we take all concentrated noise polluted regions as Pro-B regions. Thus, the extended regional voting should be able to retain the original voting selection, when $X < \frac{P_A(\aleph_w)-P_B(\aleph_w)}{1+P_A(\aleph_w)-P_B(\aleph_w)}$.

Since X and S_r is related by $X = |S_r|/N^2$, the above requirement on X can be expressed as $|S_r| < \frac{P_A(\aleph_w)-P_B(\aleph_w)}{(1+P_A(\aleph_w)-P_B(\aleph_w))\times N^2}$. This should be satisfied, when $|S_c| < \frac{n_l n_w}{(n_l+r_l-1)(n_w+r_w-1)\times r_l r_w} \times \frac{P_A(\aleph_w)-P_B(\aleph_w)}{(1+P_A(\aleph_w)-P_B(\aleph_w))\times N^2}$.

Noting that $\aleph_c = |S_c| \times \alpha$, we have proved the theorem. ■

IV. Conclusions and Examples

A. Conclusions

Figure 2 illustrates the number of noise-contaminated votes that regional and national voting can accommodate before the original Pro-A decision is reversed. We are using extended regional voting in the figure because we could see an averaged situation. Nearly equilibrium cases of $\alpha - \beta = 0.02\%$ are treated in the figure. Several typical sections of the surface in figure 2 is shown in figure 3.

We could see that, as the size of subdivided regions decreases, the number of noise contaminated votes a regional voting can accommodate increases continuously up to a certain point beyond which this starts to decrease sharply until it asymptotes to that of the national voting where the unit of voting cell and the region size are of the same order so that the improved

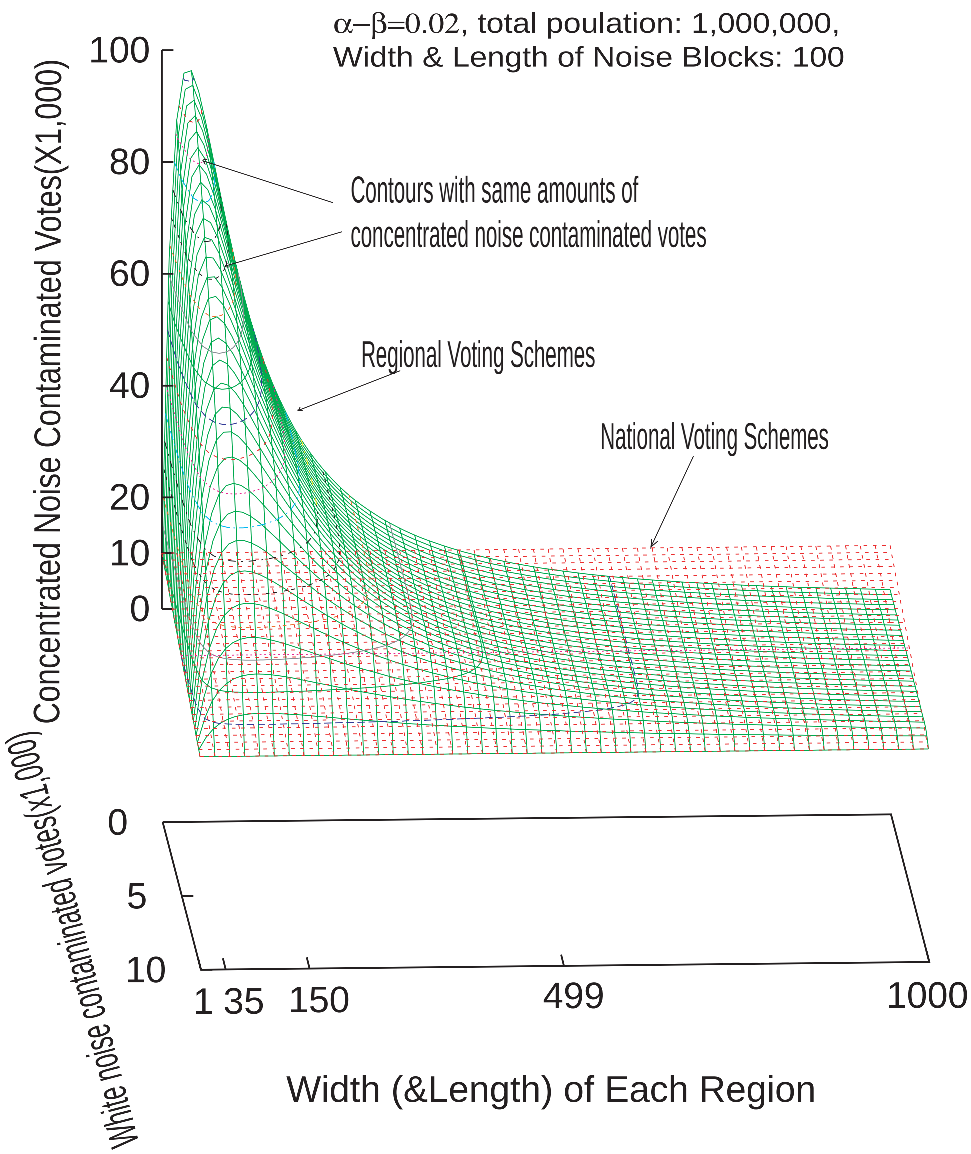


Fig. 2. Numbers of while & concentrated noise contaminated votes a voting system can accommodate

stability of the regional voting by localizing the effects of noise into a restricted number of smaller effective area may not be fully utilized in the limit. We see that the regional voting asymptotes to the national voting in both of the two extreme limiting processes; when the region size increases to that of the nation as well as when the region size decreases to the unit size of a vote

It may seem that for very large regions with small white noise, the stability margin for the

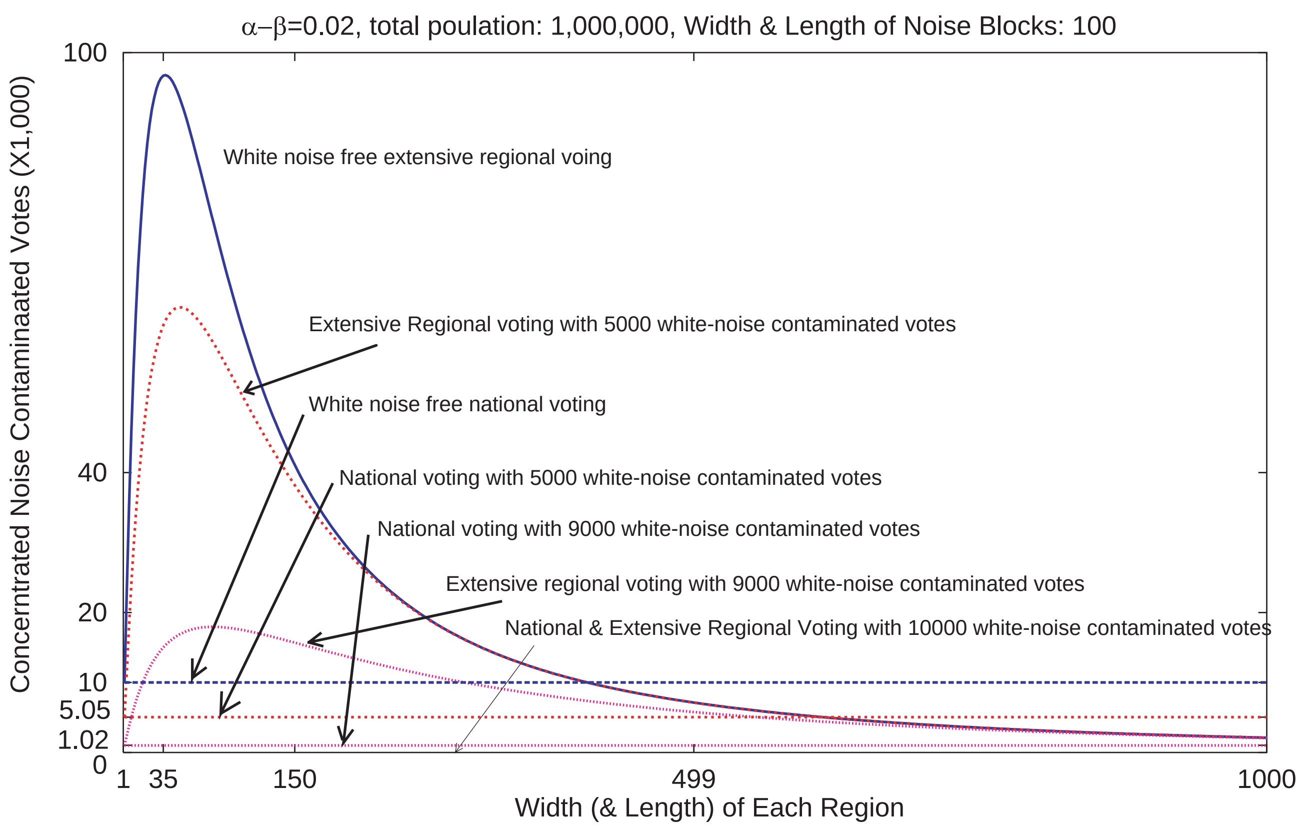


Fig. 3. Number of noise contaminated contaminated votes that a voting system can accommodate

regional voting looks smaller than that of the national voting for concentrated noise. This is due to the upper ceiling operations we have adopted in the analysis for obtaining the largest number of concentrated noise polluted regions at the worst case, where we have regarded all of the concentrated noise polluted regions as Pro-B regions in regional voting (see proofs of Theorems 2 and 3), regarding only those regions that remain entirely *free of noise contamination* to remain Pro-A. In fact, many of the Pro-B transformed regions still remain pro-*A* even in the presence of concentrated noise. Evidently this overestimation is most serious when the size of regions is large where the regional voting asymptotes to the national voting with the size of regions increasing to the size of the nation. we expect that, if the effect of over-estimation of the concentrated noise polluted regions is properly taken into account, the stability margin for the regional voting with large sized regions will increase, elevating t the surface of figure 2 representing the regional voting slightly up from the peak point to the right-back corner (consequently, the curves in figure

3 will be elevated slightly up from the peak points to the right ends). Our analysis convincingly shows that the regional voting is always more stable than the national voting, and the regional voting and the national voting will become identical, when the size of regions is so small as 1 or so large as that of the nation.

The advantage of the regional voting can be more easily understood from the physical insight: in the regional voting, we can localize the effects of concentrated noise into a restricted number of smaller regions leaving all other regions unaffected. We could see that the sets of concentrated noise polluted regions, which are the regions for surrounding the concentrated noise, are differ with different partitions. The extended regional voting actually takes an average of the numbers of noise polluted regions. It also shows that we may be able to find a best partition for the regional voting, if the distribution of concentrated noise is fixed.

In all likelihood, as our theorems show that in the absence of concentrated noise, both of national voting and regional voting are able to accommodate the same amount of white noise $\frac{1}{2}(\alpha - \beta)N$, partitioning a nation into regions does not have any influence in localizing the effect of white noise on stability. These situations are illustrated as the front bottom line, which is a joint line of the surface representing regional voting and the surface representing national voting, in figure 2. But we believe that these "pure" situations are unlikely to occur; because such a situation actually means that the noise is dispersed uniformly over the nation *without* any exception, which makes the noise sound too regular and too predictable to be "random" noise.

Thus, we claim here:

1. The regional voting is always more stable than the national voting. (So, an electoral college is always more stable than a direct popular election in selecting a president ! [9])

2. For the regional voting, the larger the number of the regions of partitions, the number of noise we can accommodate without changing the original voting results keeps increasing continuously up to a certain limit but across this point, the stability margin keeps decreasing until it asymptotes to that of national voting with the size of each region equal to 1. This limit is realized at the largest possible size of the subdivided regions.

3. For the regional voting with fixed region size, in the case that the distribution of noise-concentrated area is fixed, we can try different partitions so as to find a best partitioning scheme

[9] President George W. Bush might be happy to confirm this :-) .

that can localize the noise into a least number of polluted regions.

B. Conjecture

We believe that the above conclusion related to national and regional version of matching still remain valid even when the decision making process involves complicated matching schemes such as features extraction by PCA method. That is:

1. The regional matching is always more stable than the national matching.

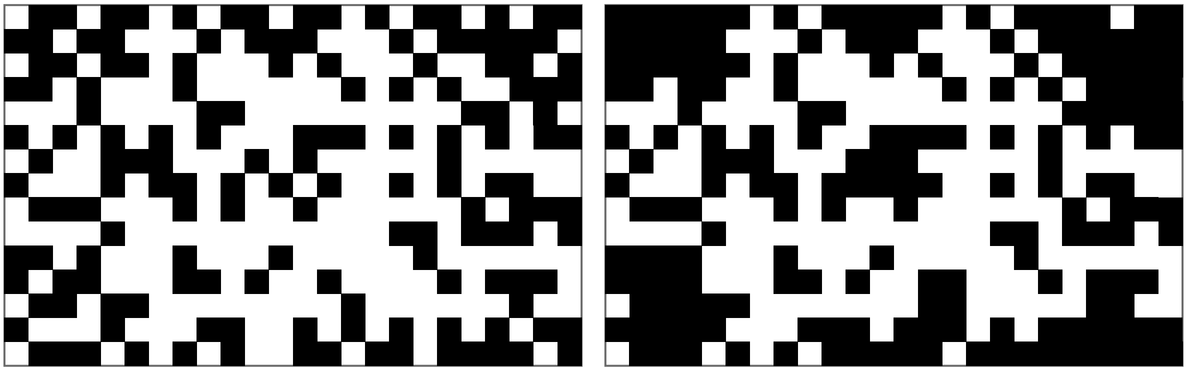
2. For the regional matching, the larger the number of the regions of partitions the more noise we are able to accommodate without changing the original voting results up to a certain limit. When the number of regions further decreases, the noise it can accommodate will decrease.

C. Experiments

1) Example 1: The first example relates to the white-black mixed flag that we have used in papers [10], [11], which we want to recognize either as a white-dominated or a black- dominated flag (see figure 4 for illustration where the cells in the figure denote a smallest unit of a "pixel"). The size of the "nation" is 24×15 cells, and the original white-dominated flag is given as in Fig. 4(a), as explained in paper [10] the noise is introduced, transforming the flag into Fig. 4(b). As it is mentioned in paper [10], if 4×5 regions are used in regional voting, the original selection of a "white" dominated flag would still remains valid, while the national voting will reverse the results of candidate selection from "white" to "black" dominated; and the 3×3 size regional voting could maintain the selection by even better margin [10].

We now try to perform the regional voting with further small regions, as indicated in Table I. A smaller regional voting of 2×3 size corresponds to situation in our theorems that the size of the regions is too small crossing a certain limit where the margin of noise accommodation starts to decrease. In this case, "White" and "Black" win the same number of regions. At the extremity, the regional voting is exactly the same as the national voting when we use only 1 cell as a nation. The result will be "White" wins 172 regions, and "Black" 188. It is surprising that the present analysis is able to explain the details of the stability diagram.

[10]In 4×5 regional voting, "white" wins 9 regions, while "black" does so in 6 with 3 regions tied; these two numbers are now 25 and 15 in 3×3 regional voting. We see $9/6 < 25/15$.



(a) (b)

Fig. 4. "White" or "Black" Dominated?

TABLE I

DOMINATING PROPERTY OF REGIONAL VOTING

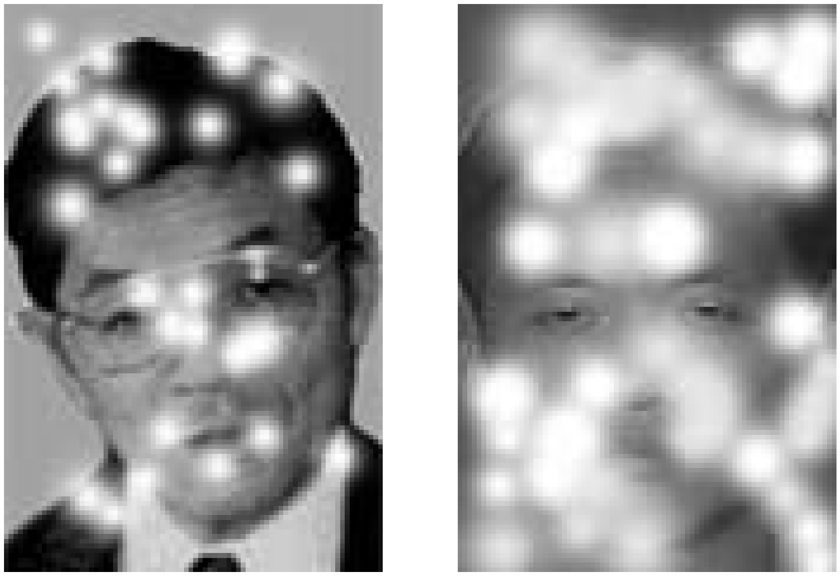
W, B and $\mathbf{W \equiv B}$ denote the numbers of "White" winning, "Black" winning and tied-up regions.

Region Size	6×5	4×5	6×3	3×3	2×3	3×1	2×1	1×1
$\mathbf{W}$	7	9	11	25	23	60	50	172
$\mathbf{B}$	5	6	7	15	23	60	58	188
$\mathbf{W \equiv B}$	0	3	2	0	14	0	72	0
$\frac{\mathbf{W}+\frac{\mathbf{W \equiv B}}{2}}{\mathbf{W}+\mathbf{B}+\mathbf{W \equiv B}}$	0.5833	0.5833	0.6	0.625	0.5	0.5	0.478	0.478

2) Example 2 : The experiments on facial recognition of 16 people's images [11] at a lower and a higher noise levels as shown in figures 5(a) and (b) have been carried out in [10]. We use noise-free images in training and these noise polluted images in testing.

The size of each picture is 80×120. As it is explained in [10], we use Turk and Pentland's PCA scheme-based eigenface approach, in the national matching scheme, and also as the way to determine the winner in each region in regional matching scheme. The recognition rates of both lower and higher noise level images with national matching and regional matching styles are demonstrated in [10], and are shown also in table II. We could see that when the size of

[11]We use images of the 16 people in Turk and Pentland's original work [14] and remove the backgrounds to get the original noise-free pictures; and independently introduce circular blocks of noise into these images to produce two sets of noise polluted test images. Fig. 5(a) and (b) are simply examples of noise polluted images.



(a)Lower Noise Level *(b)Higher Noise Level*

Fig. 5. Typical Noise Contaminated Images At Low and High Noise Level

regions decreases the recognition rate increases until up to a certain point beyond which point the recognition rate will decrease to the national voting limit.

TABLE II

MATCHING RESULT OF HUMAN FACES

Number of Correctly recognized faces	Size of Regions for Regional Voting						National Matching
	40×30	20×30	10×30	10×15	5×5	2×3	
Lower Noise Level	6	10	15	16	16	15	5
Higher Noise Level	1	1	2	2	6	2	1

V. DISCUSSION AND OPEN PROBLEMS

A. Discussion

1) Size of noise-concentrated blocks: The geometrical shapes and the sizes of the noise blocks must be so chosen to minimize possible errors between the total size of the area consisting of all the "concentrated" noise affected cells and the union of all non-overlapped noise blocks. Let us consider cutting some $n_l \times n_w$ blocks out of an area consisting of reasonably concentrated noise affected cells, or a concentrated noise affected area. It is reasonable to assume that if $n_l \times n_w$ is small compared to the size of any continuous parts of a concentrated noise affected area. We expect that a measure of what is left after cutting out blocks is surely quite small compared to the whole number of noise so that our model on concentrated noise remains valid. This is demonstrated by an important observation following two definitions.

Definition 1: A line segment is called an ortho-diameter of a continuous area of the nation, *if and only if*

1) all the points of the segment lie within the area;
2) only the two end points of the line segment lie on the boundary of the area;
3) the line segment is parallel to the horizontal line or the vertical line comprising the boundary of the nation.

Definition 2: The ortho-measure of an area (continuous or detached) of the nation is defined by the length of the shortest ortho-diameter of any continuous part of the area.

Observation 1: Let us cut out as many $n_l \times n_w$ blocks as possible and let S_c be the union of all these blocks (i.e. the noise-concentrated area). Suppose *OM* and $\aleph_n$ be the ortho-measure and the total size of the noise set that looks concentrated (all the noise except white noise). We have:

$$\lim_{\frac{OM}{\max\{n_l, n_w\}-1} \to \infty} \frac{\aleph_n - S_c}{\aleph_n} = \lim_{\frac{OM}{\max\{n_l, n_w\}-1} \to \infty} \frac{\aleph_n - S_c}{S_c} = 0.$$

Notice that the above observation includes the situation that $\aleph_n - S_c = 0$ when $n_l = n_w = 1$.

2) Extension to three or more candidates : We will show the validity of our stability properties for three or more candidates although it is true that the special coefficients of equations in the theorems 1, 2 and 3 may change depending on the definitions of noise and other conditions. This will be demonstrated by the following theorem where we have considered only 3 candidates in the presence of noise against candidate *A*.

Theorem 4: Let the percentage of the votes for candidates *A*, *B* and *P* be α, β and p respectively. Assume further that $\alpha > \beta > p$, $\alpha - \beta < p$, $\alpha + \beta + p = 1$. The anti-A-noise contaminated votes are defined as the votes that convert the votes originally for *P* to *B*. Then

$$\aleph_c < \frac{\frac{n_l n_w}{r_l r_w}}{(\lceil \frac{n_l - 1}{r_l} \rceil + 1)(\lceil \frac{n_w - 1}{r_w} \rceil + 1)} \cdot \frac{P_A(\aleph_w) - P_B(\aleph_w)}{1 + P_A(\aleph_w) - P_B(\aleph_w)} \cdot p \cdot N,$$

and

$$\aleph_w < (\alpha - \beta) \times N$$

is a sufficient condition for regional voting to preserve the original candidate selection, where $P_A(\aleph_w)$ and $P_B(\aleph_w)$ denote the percentages of A-pro and B-pro regions in the presence of white noise only and can be calculated as

$P_A(\aleph_w) = \sum_{x,y \leq z}^{x+y+z=r_l r_w} \frac{(r_l r_w)!}{x!y!z!} (p - \aleph_w/N)^x \alpha^y (\beta + \aleph_w/N)^z)$

$P_B(\aleph_w) = \sum_{x,z \leq y}^{x+y+z=r_l r_w} \frac{(r_l r_w)!}{x!y!z!} (p - \aleph_w/N)^x \alpha^y (\beta + \aleph_w/N)^z)$

The sufficient and necessary condition for national voting to preserve the candidate selection is $\aleph_c + \aleph_w \times \frac{N - \aleph_c/p}{N} < (\alpha - \beta)N$.

Theorem 4 clearly confirms that the regional voting can still accommodate more noise contaminated votes than national voting when α and β are very close.

3) Parameter Estimation: Theorems 2 and 3 show the feature of stability for regional voting with respect to the size of the regions as demonstrated in Fig. 3. From application point of view, of course, we would like to choose the size of the regions in such a way as to accommodate the largest amount of noise-contaminated votes, which critically depend on important physical parameters such as the level of white noise, size of noise block etc. To do this, we should have good knowledge of the parameters involved in the formulas. However, these parameters are usually application and environment dependent which we believe are difficult to analyze and are beyond the main aim of this paper. What we can teach here is, when we have no knowledge of these parameters, we have to try different sized partitions to find the best region size in setting up regional matching scheme for machine recognition under a given environment. The unimodality of the stability margins shown in the curves in Fig. 3 and also demonstrated in Table II seems to suggest that a revised binary searching (Dichotomous Searching) approach [15] is possible in finding such an optimal region size for regional voting purpose.

B. Further Work and Open Problems

The concept of regions in regional voting need not be restricted to geometrical interpretation. A region can be considered as a society of people having same ethnic backgrounds or sharing same interest in political area, or as a set of information of same frequency or having same characteristic features in speech recognition and image processing. A regional voting with partitioned regions using these criteria shall also retain its stability advantage. The applications using such a regional voting scheme are of great interest, and may become new topics as traditional voting theory does not consider an issue of stability.

Some interesting voting schemes may emerge from our regional voting scheme to extend the applicability of the method to a wider range of decision making processes. For example, it seems to be a highly exciting subject to pursue the analysis of a multi-level regional scheme where we employ the regional voting scheme recursively for determining the winners of *regions* by recursively partitioning each of the regions into smaller (sub)regions.

The stability analysis of other kinds of voting, such as the weighted-voting system, and a variety of voting systems in political science, in the presence of different kinds of noise, is also of great interest.

REFERENCES

[1] Sunzi, Lionel Giles, Sun Tzu "The Art of War," Dover Publication, November, 2002.

[2] Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349–1380, 2000.

[3] R. W. Picard and T. P. Minka, "Vision texure for annotation," *Multimedia Systems*, vol. 3, pp. 3–14, 1995.

[4] Daniel Keren, "Recognizing image "style" and activities in video using local features and naive bayes," *Pattern Recognition Letters*, vol. 24, no. 16, pp. 2913–2922, 2003.

[5] K. J. Arrow, *Social Choice and Individual Values,* Yale University Press, 2nd edition, 1963.

[6] R. D. Luce and H. Raiffa, *Games and Decisions – Introduction and Critical Survery*, John Wiley and Sons, 1957.

[7] M. A. Garcı́a, S. Velázuez, and A. D. Sappa, "A two-stage algorithm for planning the next view from range images," in *Proceedings of the 9th British Machine Vision Conference*, John N. Carter and Mark S. Nixon, Eds., Southampton UK, September 14-17 1998, pp. 720–729.

[8] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 530–535, 1997.

[9] J. J. Bloch, D. S. Daniels, and A. Z. Spector, "A weighted voting algorithm for replicated directories," *Journal of the ACM*, vol. 34, no. 4, pp. 859–909, October 1987.

[10] L. Chen and N. Tokuda, "Robustness of regional matching scheme over globe matching scheme," *Artificial Intelligence*, vol. 144, no. 1-2, pp. 213–232, 2003.

[11] L. Chen and N. Tokuda, "Stability analysis of regional and national voting schemes by a continuous model," *IEEE Transactions on Knowledge and Data Engineering*, vol. 15, no. 4, pp. 1037–1042, 2003.

[12] L. F. Cranor, *Declared-Strategy Voting: An Instrument for Group Decision-Making*, Ph.D. thesis, Department of Engineering and Policy, Washington University, December 1996.

[13] Jay B. West and J. Michael Fitzpatrick, “The distribution of target registration error in rigid-body, point-based registration,” *IEEE Transactions on Medical Imaging*, vol. 20, no. 9, pp. 917–927, 2001.

[14] M. Turk and A. Pentland, “Eigenfaces for recognition,” *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[15] G. Brassard and P. Bratley, *Fundamentals of Algorithmics*, Prentice-Hall, Inc., 1996, p.226.