

ART 모델의 multirun 횟수 감소에 관한 연구

김미나, 김도년, 조동섭
이화여자대학교 전자계산학과

A Study on decreasing the Number of Multirun in ART Model

Kim Mi-Na, Kim Do-Nyun, Cho Dong-Sub
Dept. of Computer Science, Ewha Womans University

Abstract

The ART(Adaptive Resonance Theory) model is self-organized with nonstationary input patterns in real time. But there is a multirun problem caused by fault clustering, or pertubated clustering and confines the advantage of the stationary real-time processing in ART model. In this paper, we propose the incremental vigilance threshold approach to decrease the number of multiruns. The incremental vigilance threshold approach is to learn with incremental vigilance threshold and competition with clusters.

1. 서론

ART(Adaptive Resonance Theory) 모델은 불안정한 입력패턴으로부터 실시간에 안정된 인식으로 자기조직화한다. 그러나 ART 모델의 이러한 실시간 처리의 장점을 제한하는 요소로써 multirun 문제가 있다. 즉, 비슷한 입력패턴일지라도 발생순서에 따라서 분류되는 클러스터값이 달라지게 된다. 따라서 이러한 문제를 해결하기 위하여 입력 패턴의 그룹을 여러번 실행해야 하는 문제가 발생한다.

본 논문에서는 vigilance 임계치를 가변적으로 바꾸어 주고 또한 여러번 수행되는 주기안에서 클러스터들끼리 경쟁을 시키는 증가 vigilance 임계변수 방법을 사용해서 multirun의 횟수를 줄이기 위한 방법을 제안하고 실험해 보고자 한다.

2. ART 모델

신경망 모델은 훈련 방법에 따라서, 목표값(target value)을 함께 학습시키는 지도 학습(supervised learning) 방법과 목표값 없이 자기조직화를 이루는 자율학습(unsupervised learning)으로 나눌 수 있다. ART 모델은 자율학습 신경망으로써 코호넨의 형상지도 모델이 가진 경쟁학습의 약점인 안정성으로 보강한 모델이다. ART 모델의 가장 큰 특징은 기존에 학습되었던 정보들이 새로운 학습에 의해서 지워지지 않도록 하면서, 새로운 지식을 자동적으로 전체 지식베이스에 일관성 있는 방법으로 통합한다는 것이다. 즉, 적절하게 대응되는 새로운 정보를 이용하여 이미 배운 내용을 정제(refine)하며, 새로운 지식 카테고리의 학습을 위하여 새로운 노드를 생성하고, 기억용량을 넘어서는 과도한 새로운 입력때문에 기존에 취득한 내용이 지워지는 것을 방지한다. 따라서 끊임없이 변하는 환경에서 자신의 메모리 용량을 전부 소모할 때까지 제한없는 입력에 대해 실시간으로 빠르고 안정되게 학습시킬 수 있는 구조이다.[3]

Carpenter와 Grossberg에 의해 개발된 ART 모델은 이진수로 된 입력만을 받아들일 수 있는 ART1 모델이 있다. 그러나 ART1 모델은 이진수로 표현된 패턴만을 입력으로 받아들일 수 있으므로 별도의 전처리 과정이 필요하다. 또한 저장방법이 비효율적이어서 많은 수의 대표패턴을 저장할 수 없으며 실제계 구현시에 대부분 하드웨어적인 지원이 요구된다. ART2 모델은 이진 패턴뿐만 아니라 아날로그나 계조영상의 벡터 컴포넌트도 입력 패턴으로 받아들일 수 있는 ART 모델이다. 최근에 들어 화학전달물질을 ART 네트워크에 결합한 ART3 모델도 연구되었다.

2.1. ART2 모델의 구조

ART2 모델은 두개의 하부구조로서 주의 서브시스템과 적응 서브시스템으로 구성되어 있으며, 주의 서브시스템에는 F1, F2, 두개 계층이 있어, 각 계층은 노드들로 구성되어 있으며, 계층의 노드들은 다른 계층과 완전히 연결되어 있다. 다른 구성 요소들간의 연결은 화살표로 나타내며, +표시가 있는 화살표는 흥분 신호이며, -로 표시되어 있는 화살표는 억제 신호를 나타낸다. ART2 모델의 시스템 구조도는 그림1.과 같다.[2]

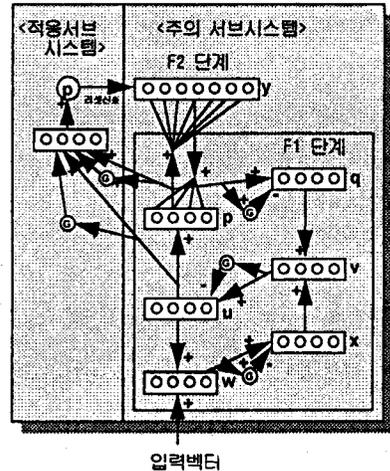


그림 1. ART2 시스템의 구조도

2.2. ART2 모델의 알고리즘[1]

x 와 W_j 는 각각 입력 벡터와 뉴런 j 의 가중치라 하며, 승자를 정하는 기준은 Euclidian 거리나 다른 거리 공식에 의해서 최소 거리를 구함으로써 선택한다.

- ① 학습패턴이 주어져 있을 때, MINNET(Minimum Net)을 적용시켜서 최소거리 $\|x - W_j\|$ 를 가지는 승자 j^* 를 정한다.
- ② $Vigilance Test: \|x - W_j\| < \rho$ 를 만족하면, vigilance test를 통과했다고 하며, vigilance 값 ρ 는 클러스터의 반지름이다.
- ③ 승자가 vigilance 테스트에서 실패하면, 가중치로서 $W_k = x$ 를 가지는 새로운 뉴런 k 를 생성한다.
- ④ 승자가 vigilance 테스트에 성공하면, 승자 j^* 의 가중치를 다음처럼 갱신한다.

$$W_j^{(new)} = \frac{x + W_j^{(old)} |cluster_j^{(old)}|}{|cluster_j^{(old)}| + 1}$$

여기에서 $|cluster_j|$ 는 $cluster_j$ 에 있는 구성원의 수를 나타낸다.

3. ART2 모델의 문제점 : multirun 문제

3.1. multirun 문제[1]

ART2 모델에 있어서 문제점을 S. Y. Kung이 다음과 같이 제시했다 [1]. ART2 모델은 vigilance 임계치와 입력 패턴의 순서에 따라서, 각 클러스터에 속하는 패턴들이 달라질 수 있고, 클러스터의 중심값이 다르거나 클러스터수가 다른 클러스터링의 결과가 생길 수 있다.

- ① 입력패턴의 입력 순서에 따라서 생기는 결과 클러스터의 수가 달라질 수 있다. 다음의 예는 10개의 패턴 2개의 속성을 가지는 입력 패턴으로써 단지 입력 순서가 다른 패턴에 대해서 구한 결과이다. 예에서 보면 두 실험의 결과로써 생기는 클러스터의 수가 다를 수 있다. 즉, 실험(a)는 3개의 클러스터를 가지며, 실험(b)는 2개의 클러스터를 가지게 된다.

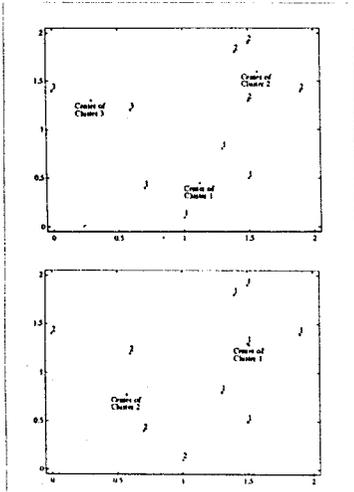


그림 2. (a) ART2의 학습 결과. (b) 클러스터링을 다시 시킨 결과. 그림에서 나타나는 선은 클러스터 중심값에 대한 경계선이다.

- ② vigilance 값의 크기에 따라서 결과가 다르게 나타날 수 있다. 일반적으로 vigilance 임계치가 작을수록 더 많은 클러스터가 생기는 경향을 갖는다.
 ③ 결과적으로 클러스터링을 한 후에 조사해 보면, 초기에 입력된 패턴들이 다른 클러스터에 속해 있는 경우가 있다. 그림에서 보면 (a)에서 실제로 클러스터 3에 속해야 하지만, ART2가 1, 2로 잘못 분류한 패턴이 나타나고 있다. (b)는 다시 클러스터링을 한 결과이다.
 ④의 문제점을 해결하기 위해서 multirun을 해야 한다. 즉, 잘못된 분류가 고쳐질 때까지 즉, 더 이상 클러스터의 구성원이나 중심값이 달라지지 않을 때까지 같은 패턴 집합에 대해서 여러번 ART 학습을 시키게 된다.

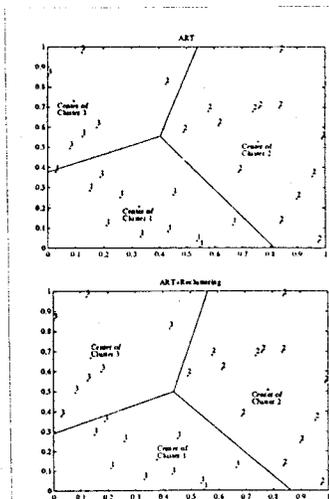


그림 3. 입력패턴의 순서에 따라서 나타난 다른 결과. (a) 원래의 순서인 경우 (b) 역순인 경우

3.2. 제안 방법

본 논문에서는 vigilance 값을 변화시키면서 multirun을 수행하는 두가지 방법, 증가 임계치(incremental vigilance threshold) 방법과 감소 임계치(decremental vigilance threshold) 방법을 제시하고, 증가 임계치 방법, 감소 임계치 방법에 대해서 실험해 보았다. 증가 임계치 방법과 감소 임계치 방법 차이는 시작하는 초기의 vigilance 임계치 ρ_0 와 변하는 delta의 방향이 다르며, m번의 multirun을 실행한 후에 최종 vigilance 임계치 ρ_m 는 같게 된다. vigilance 임계치 ρ_0 와 multirun의 횟수 m이 주어지고, i번째의 run을 실행할 때, 다음의 vigilance 임계치를 사용해서 ART2 신경망으로 학습을 시키게 된다. 감소 임계치 방법 일 때 주어지는 vigilance 임계치는 다음과 같다.

$$\delta = \frac{Ratio * \rho_0}{m}$$

$$\rho_i = (1 + Ratio) * \rho_0 - \delta \text{ for } i=1,2,\dots,m$$

3.2.1. 증가 임계치 방법

두번째 방법으로써 증가 임계치 방법은 i번째 run에서 ρ_i 는 다음의 식에 의해서 vigilance 임계치를 가지고 수행하게 된다.

$$\delta = \frac{Ratio * \rho_0}{m}$$

$$\rho_i = (1 - Ratio) * \rho_0 + \delta \text{ for } i=1,2,\dots,m$$

Ratio는 증분의 크기 비율을 나타내는 상수이며, ρ_0 은 학습 초기에 주어지는 vigilance 임계치이다.

그러나 증가 임계치 방법만을 사용해서 작은 vigilance 임계치 ρ_i 로 학습한 후에 δ 만큼 증가한 vigilance 임계치 ρ_{i+1} 로 학습을 하면, 더 이상의 클러스터의 변이가 일어나지 않고 많은 수의 클러스터를 가진 상태로 고정되어므로, 고정 크기의 ρ_0 를 가지는 실험 결과에 비해서 좋은 결과를 얻지 못했다. 너무 많이 생기는 클러스터의 수를 줄이기 위해서, vigilance 임계치 ρ_i 에서 분류된 클러스터들을 다음번 학습을 하기 전에 클러스터들간의 경쟁 알고리즘을 사용해서 경쟁에서 이긴 클러스터들에 대해서만 다음의 vigilance 임계치 ρ_{i+1} 을 가지고 학습을 시키는 방법을 사용하였다.

3.2.2. 클러스터 경쟁 알고리즘

i번째 run을 수행한 후에, 다음 방법에 의해서 클러스터들을 병합한다.

for all cluster, and $j < k$

$$\min_distance_j = \min \left\{ \frac{|cluster_j| * b_j + |cluster_k| * b_k}{|cluster_j| + |cluster_k|} \right\}$$

if ($\min_distance_j < \rho_{i+1}$)

$$b_k = \frac{|cluster_j| * b_j + |cluster_k| * b_k}{|cluster_j| + |cluster_k|}$$

$$|cluster_k| = |cluster_j| + |cluster_k|$$

delete(b_j)

$$|cluster| = |cluster| - 1$$

여기서 b_j 는 cluster_j의 중심값이 되고, $|cluster_j|$ 는 cluster_j에 속하는 패턴의 수이며, $|cluster|$ 는 전체 클러스터의 수이다.

3.2.3. 증가 임계치 방법의 과정

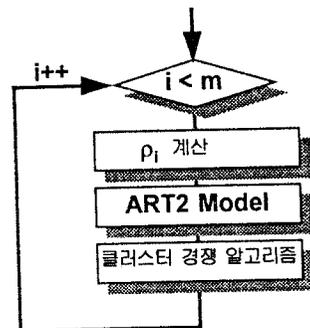


그림 4. 증가 임계치 방법의 실행 과정

4. 실험 및 고찰

실험 결과를 비교해 보기 위해서 대조실험으로써 상수 vigilance 임계치를 사용해서 m 번의 run을 수행하는 동안 임계치는 변경시키지 않는 상수임계치 방법에 대해서도 동일한 실험을 하였다. 실험결과를 보면 클러스터의 갯수와 각 클러스터에 속하는 구성원들이 약간씩 다르므로 좋고 나쁨을 판별하기 위한 기준으로 비용함수(energy function)을 다음과 같이 정의한다.

$$E = \sum \|b_i - \text{pattern}_i\|$$

for all $\text{pattern}_i \in \text{cluster}_i$

여기에서 b_i 는 cluster_i 의 중심값이다.

실험 대상은 2개의 영상, (a), (b)를 사용하였다. 영상(b)는 문자가 들어간 혼재문서 영상이다. 아래의 영상에서 각각 크기 100×32 의 서브영상 4개를 임의로 정하고 크기 4×4 의 볼록으로 나누어서 200개의 입력패턴으로 사용하였다. Pao의 ART model을 참고해서 여기에 증가, 감소 vigilance 임계치 계산과 클러스터들 끼리의 경쟁에 적합하도록 코딩했다[4].

실험 결과는 표1.~표2에 있다. 표에서 나타나는 CVT는 상수임계치(constant vigilance threshold) 방법을 말하며, DVT는 감소임계치(decremental vigilance threshold) 방법을, IVT는 증가임계치(incremental vigilance threshold) 방법을 말한다. 또한 vigilance 임계치 ρ 는 5번 run후에 갖게 되는 목표임계치이다. 상수임계치 방법인 경우에는 초기임계치와 목표임계치가 같다.



그림 5. 전체영상 (a), (b)

5. 결론

같은 학습 데이터에 대해서 multirun을 수행하는 동안에, 상수임계치 방법, 감소임계치 방법, 증가임계치 방법을 사용해서 학습데이터를 분류해보고 결과적으로 생긴 클러스터들의 에너지 함수값을 구해 보았다. 실험 결과로써 클러스터들끼리의 경쟁을 사용한 증가임계치 방법이 vigilance 임계치에 따라서 다소 차이는 있으나 대체로 작은 에너지 함수값을 가지는 경향이 있었다.

향후 연구 과제로서 상수임계치 방법, 감소임계치 방법, 증가임계치 방법에 대한 다각적인 실험방법과 객관적인 평가요소를 정하고 그에 따른 평가를 할 필요가 있다. 따라서 분류된 패턴을 이용한 영상복원이나 인식 실험을 통해서 실험 결과를 검증할 필요가 있다. 여기에 덧붙여서 vigilance의 증분의 비율의 설정에 따른 변화에 대해서 실험해 보고, 증분을 선형으로 주는 방법외에 다른 함수를 사용할 때의 변화를 비교할 필요가 있다. 또한 좋은 성능을 가지는 ART모델의 vigilance 임계치와 multirun의 횟수와의 연관성을 알아보고 최적의 값을 얻기 위한 연구가 필요하다.

참고문헌

- [1] S.Y.Kung, Digital Neural Networks, pp.79-85, Prentice Hall, 1993.
- [2] James A. Freeman, David M. Skapura, Neural Networks algorithms, applications, and programming techniques, pp.291-298,316-336, Addison Wesley, 1991.
- [3] 김갑주, "ART를 이용한 혼재문서의 이진화", 석사학위논문, 1994.
- [4] Yoh-Han Pao, Adaptive Pattern Recognition and Neural Networks, pp.171-182, 291-299, Addison Wesley, 1989.

표 1. 영상(a)의 서브영상, DATA1, PIN, DATA2, PIN, DATA3, PIN, DATA4, PIN파일에 대한 실험결과

vigilance threshold ρ	CVT	DVT	IVT	CVT	DVT	IVT	CVT	DVT	IVT	CVT	DVT	IVT
9.5000	123.1833	117.4569	121.4995	63.1088	63.1088	79.2010	80.0145	77.2884	79.2645	88.8719	93.1183	95.3492
9.0000	123.9197	117.4569	124.6768	72.0715	63.1088	75.5470	80.0145	77.2884	77.7286	95.5142	85.9107	95.3492
8.5000	125.0069	121.7598	121.3196	72.0715	63.1088	81.9357	80.0145	82.0989	78.7426	94.5964	85.9107	100.6823
8.0000	122.5567	120.2682	118.9718	75.9071	63.1088	84.8078	78.6154	80.0145	74.9179	97.2128	85.9107	102.6407
7.5000	124.0181	123.1833	102.1068	82.7563	63.1088	82.0678	75.5807	80.0145	72.7912	97.5684	92.1284	101.6951
7.0000	115.9467	123.9197	82.5049	82.7563	72.0715	77.3816	75.5807	80.0145	64.6793	97.5684	94.5964	106.7335
6.5000	93.4206	122.2947	70.5846	85.3306	76.6379	75.6022	72.4110	78.6154	63.3473	103.1106	97.2128	98.8861
6.0000	84.6198	120.9585	70.1410	79.8908	82.7563	54.6071	71.2885	75.5807	55.4861	106.7895	97.5684	104.1990
5.5000	72.7951	113.3915	61.8300	74.8743	82.7563	46.7291	67.6261	72.4110	58.5349	105.1097	102.1141	98.8574
5.0000	69.4743	88.9019	53.5205	56.7695	82.4081	51.0290	57.0707	72.4110	51.8530	106.1097	101.5475	90.2215
4.5000	55.6009	75.8131	47.4028	54.2905	75.1783	42.4576	53.1996	67.8327	42.6505	94.1639	102.8562	93.7013
4.0000	51.1564	68.9372	40.0842	45.3027	57.3248	36.4306	52.6471	67.2797	34.2251	95.5363	103.7541	82.4003
3.5000	43.5237	54.2885	35.3852	41.0563	53.3708	25.6470	34.6660	53.4088	33.6164	86.0444	94.0406	65.3948

표 2. 영상(b)의 서브영상, DATA1, PIN, DATA2, PIN, DATA3, PIN, DATA4, PIN파일에 대한 실험결과

vigilance threshold ρ	CVT	DVT	IVT	CVT	DVT	IVT	CVT	DVT	IVT	CVT	DVT	IVT
9.5000	126.4692	131.0480	111.1043	168.9974	218.5108	157.2199	183.0007	207.2337	151.2880	85.9094	109.7184	80.1659
9.0000	109.5692	128.1219	108.1263	166.2354	194.5845	149.8663	183.5715	213.2531	136.4936	85.9094	101.5420	73.5149
8.5000	109.5692	128.1219	107.4137	166.2354	189.9234	130.2693	173.3785	203.9845	119.6490	85.9094	100.2682	70.1246
8.0000	121.5541	125.0867	107.4128	159.4166	177.2669	121.5973	157.8681	188.2833	119.6490	80.1659	100.3660	71.2683
7.5000	114.9078	118.3760	107.3353	135.4885	168.9974	109.3215	130.4299	187.9631	110.5547	73.5149	85.9094	71.2683
7.0000	114.1722	109.5692	109.2642	128.7016	166.2354	111.0195	119.6366	176.8905	96.3587	71.2683	85.9094	68.3318
6.5000	112.7485	116.3473	110.6604	111.1890	159.4166	98.4409	122.3564	158.4383	84.7617	71.2683	80.1659	68.1995
6.0000	106.6760	114.9078	100.1228	111.0195	135.4885	95.8052	104.0956	127.3754	78.9245	68.3318	73.5149	66.2424
5.5000	104.0512	111.7200	96.9128	98.4409	125.8777	82.7247	85.0967	119.6366	75.2175	65.3039	71.2683	64.1956
5.0000	101.3652	106.6760	91.2983	94.7395	111.0195	75.6485	77.4529	108.3944	65.0600	64.4014	71.2683	63.4991
4.5000	93.7337	105.0867	86.1213	86.4938	100.7229	57.6265	74.0026	85.0967	59.4820	64.7597	65.3039	66.4830
4.0000	91.2147	98.5599	80.8965	70.1855	94.7395	52.9142	66.6153	77.4529	56.7595	64.7597	64.4014	68.7246
3.5000	82.3882	91.2711	75.1831	53.4380	83.3591	46.6660	59.1105	71.9805	54.8878	67.3572	64.7597	67.8513