

**ИНФОРМАЦИОННИ АСПЕКТИ НА ПРОЦЕСИТЕ ПРИ ОБЕКТИВНО
ОПРЕДЕЛЯНЕ КАЧЕСТВОТО НА ХРАНИ****INFORMATION ASPECTS OF THE PROCESSES DURING AN OBJECTIVE
DETERMINATION THE QUALITY OF FOODS****Лена Костадинова***Университет по хранителни технологии***Атанас Георгиев***Университет по хранителни технологии***Николай Шопов***Университет по хранителни технологии***Abstract**

The results from analyzing the processes when determination the quality of food according to the developed more than 60 years ago Shannon's theory are presented in the paper. On the base of the published in 1948 fundamental work "A mathematical theory of communication" and the evolution of its treatments the scientific team formulated a new vision to one of the everlasting activity of the people insuring their food. All of the more or less important achievements of the civilization continuously are accompanied by the development of the science and the practice of food. This determines as a first priority of the mankind such activity, concerning food, without which is impossible not only qualitative live, but also the existence on the Earth itself. In the paper an attempt has been made for adapting the basic formulations from the theory of information to the task for improving from a strait scientific positions of the vision on the essence of the processes in an objective (machine) determination of the quality of food during their production, storage, transportation and consuming, i.e. from the "farm to the fork". In its deep essence this chain is not only material-energetic, but also to the large extend – information i.e. it is a combination of material-energetic and information systems (MES and IS. This paper presents the results from the analysis and improvement of the processes in the second type systems

Keywords: information, information channel, entropy, amount of information, value of information, feature extraction, quality of food, traceability of food.

ВЪВЕДЕНИЕ

Едно от най-важните направления на добиването и производството на храни от растителни и животински продукти е осигуряването на възможност за обективно (машинно) определяне на тяхното качество, т.е. независимо от влиянието на човешкия фактор (при евристичния подход). При това тази възможност трябва да позволява процесите да се извършват не само в статика (лабораторно), но и в динамика (в поток), в реално време (в хода на производствения процес, по време на транспортиране, съхранение, в търговската мрежа и т. н., до достигането при консуматора). Тези идеи придобиват особена значимост с попявата на възможностите, предлагани от бързото навлизане на информационните технологии в технологичните и производствени процеси

и в по-нетрадиционните за приложението им сфери – аграрния сектор и хранителната индустрия. В 70-те години на миналия век започна бързо навлизане в практиката на машини за определяне качеството на аграрни продукти, основно плодове и зеленчуци. Процесите се извършваха автоматично, а системите се изграждаха с използване на дискретни електронни елементи. Този период не беше белязан все още от навлизането на технологично ориентирани контролери и компютърни системи. Независимо от това, конвенционалните системи за автоматично определяне на качеството на домати, праскови, кайсии, пипер, моркови, картофи и т. н. Бързо навлязоха в практиката и позволиха значителен напредък в повишаване качеството на хранителните продукти. В последните години на XX век, с бързото

развитие на компютърната техника и компютъризацията на най-важните процеси в микроелектрониката, машиностроенето, медицинската диагностика и др., започнаха и бързи процеси на компютъризация и в хранителните производства [3,4]. Това определи и новите направления в подходите на проследяемост на качеството на храните не само в рамките на едно производство, град или държава, но и в трансконтиненталните ветиги за производство и доставка на храни [2]. Това насочи изследванията към информационните аспекти на процесите за обективно определяне качеството на храните (между две произволни точки на планетата).

ИЗЛОЖЕНИЕ

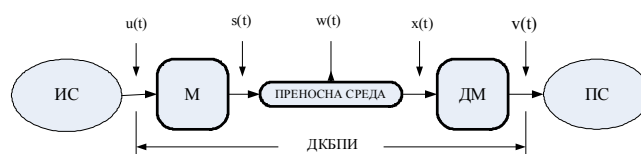
Обобщеният функционален модел на една класическа система за обективно определяне на качеството на аграрни продукти може да се представи със структурата от фиг.1.



Фиг. 1. Функционален модел на система за обективно определяне качеството на аграрни продукти

В нея са включени функционални модули за осъществяване на следните базови функции: "възприемане" на образа на продукта, намиращ се в полето на зрение на сензорния модул (СМ) (най-често на ниво "зрителни образи т.е. модул за машинно зрение"); "изграждане" на базата на образа, на "вторично описание" (изображение) на продукта, във вид на електрически сигнал, съдържащ информацията за образа (МИП); "представяне" по подходящ начин, "съхранение" и "същинска обработка" на изображението в съответствие с разпознаващия алгоритъм на използвания метод за разпознаване на образи (класификация) (МСОИ); изработване на решение в съответствие с решаващото правило на метода за разпознаване; "пространствено отделяне" (сепариране) на продукта в съответствие с решението (МИРС). Обобщената по този начин функционална структура на системата за класификация по качество (СКК) на продукта

позволява трансформирането и, в терминологията на информационните системи като "дискретен канал без памет и без изтриване" [1](фиг.2).



Фиг. 2. СКК като дискретен канал без памет и без изтриване

Тук с ИС е означен "източника на съобщение (информация) $u(t)$ " в една информационна система, а с ПС – този, за когото е предназначена информацията на ИС, т.е. за „получателя на съобщението“. За СКК това е ползвателя на крайното решение от класификацията на продукта, или сепариращото устройство. С (М) е означен условно „модулятор“, който предоставя на образа $u(t)$ неговия материален носител, във вид на сигнала $s(t)$. Последният е физична величина, съгласяваща се с преносната среда, в която $s(t)$ се разпространява. Към нея е означено и смущението $w(t)$, чието влияние придава случаен характер на информационните процеси в съобщителния канал ДКБПИ – „дискретен канал без памет и изтриване“. Модулът „демодулятор“(ДМ) представя в понятията на информационната система функциите на модулите за същинска обработка на изображението, алгоритмите за разпознаване и изработване на крайното решение за класификация в разпознаващата система от фиг.1. тук с $x(t)$ и $v(t)$ са представени съответно модифицираните (изкривени) варианти на сигнала $s(t)$ и съобщението $u(t)$, като следствие от случайния характер на процесите в системата, породени от $w(t)$ – за информационната система или от функционирането в условията на непълна информация за образа на обекта – в системата за разпознаване на образи.

Трансформираната функционална структура има своето аналитично и графично описание в общата теория на сигналите, респективно теорията на информационните системи [1].

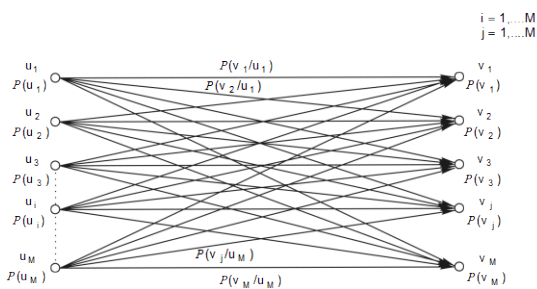
За дискретния канал без памет и без изтриване е в сила равенството:

$$P(v_j) = \sum_{i=1}^M P(u_i) P(v_j/u_i), \quad i=1, 2, \dots, M; \quad (1)$$

$$j=1, 2, \dots, M,$$

$$[P(v_j/u_i)] = \begin{bmatrix} P(v_1/u_1) & P(v_2/u_1) & \dots & P(v_j/u_1) & \dots & P(v_M/u_1) \\ P(v_1/u_2) & P(v_2/u_2) & \dots & P(v_j/u_2) & \dots & P(v_M/u_2) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ P(v_1/u_i) & P(v_2/u_i) & \dots & P(v_j/u_i) & \dots & P(v_M/u_i) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ P(v_1/u_M) & P(v_2/u_M) & \dots & P(v_j/u_M) & \dots & P(v_M/u_M) \end{bmatrix} \quad (2)$$

И графите на условната и безусловна вероятности:



Фиг. 4.2. Модел на дискретен канал без памет и без изтриване

Фиг. 3. Дискретен канал без памет и без изтриване

С използване на мярката на Шенон за определяне на количеството информация може да бъде формирана оценката на пренесеното по канала и на изгубеното в него, количество информация. С използване на модела се определят неизвестните вероятностни характеристики

$$P(u_i, v_j) = P(u_i)P(v_j/u_i) \quad \text{и} \quad P(v_j) = \sum_{i=1}^M P(v_j, u_i), \quad (3)$$

а от там и изхода на канала

$$P(u_i/v_j) = P(u_i, v_j)/P(v_j). \quad (4)$$

Ако, например, е прието на изхода съобщението v_j , остава известна неопределеност относно действителния, реализиран при това, преход, т.е. $P(u/v) = P(?/v)$. Необходимо е, по Шенон, количество информация за компенсиране на неопределеността, представлява усредненото по всички съобщения u_i количества $I = \log \{1/P(u_i/v_j)\}$:

$$- \sum_i P(u_i/v_j) \log P(u_i/v_j), \quad (5)$$

а за всички възможни двойки съобщения (u_i, v_j) от входния и изходния, ансамбли $\{u_i\}$ и $\{v_j\}$, $i=1, 2, \dots, M$ и $j=1, 2, \dots, M$ – чрез усредняване на (5) по всички j :

$$H(U/V) = - \sum_i P(v_j) \{ \sum_i P(u_i/v_j) \log P(u_i/v_j) \}$$

или

$$H(U/V) = - \sum_i \sum_j P(u_i, v_j) \log P(u_i/v_j) \quad (6)$$

Разликата между ентропиите $H(U)$ и $H(U/V)$, дава точно реалното количество информация I_p , усреднено по всички двойки съобщения “предадено-прието”, за един сеанс на изхода на канала:

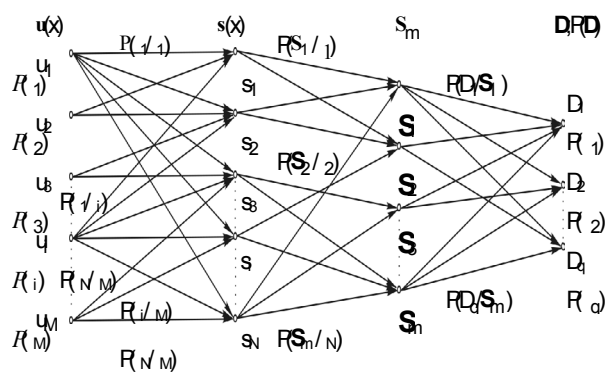
$$I_p = H(U) - H(U/V) = - \sum_{i=1}^M P(u_i) \log P(u_i) - \sum_{i=1}^M \sum_{j=1}^M P(v_j) [P(u_i/v_j) \log P(u_i/v_j)] \quad (7)$$

От теорията на информацията е известно, че основното предназначение на канала е пренасяне, със зададена скорост и достоверност, на всички съобщения на ИС (входа на канала) до ПС (изхода на канала). Това означава, че за каналите без памет и без изтриване е в сила условието $i=j=1, 2, 3, \dots, M$, т.е. равенството на обемите на входния и изходния ансамбли. (всяко съобщение, подадено на входа, трябва да бъде получено на изхода). Приемането, априори, на съотношение $i > j$, е недопустимо, то е в противоречие с казаното. Обратното неравенство $i < j$ е допустимо за разглеждане единствено за каналите без памет с изтриване, доколкото при тях в изходния ансамбъл $\{v_j\}$ се въвежда допълнителна позиция θ за “изтриване” (за съобщения, които е невъзможно да бъдат “възстановени” за ПС, т.е. за $i=1, 2, 3, \dots, M$ е допустимо $j=1, 2, 3, \dots, M, M+1$).

Като се отчитат горните постулати на информационната теория може да се моделират процесите на обективно определяне

на качеството и сортиране на аграрни продукти (изобразени чрез функционалната структура на фиг. 1).

Обобщеният модел е представен на фиг.4. Той може правомерно да се разглежда като модел на дискретен съобщителен канал без памет и без изтриване, с ретранслация.



Фиг. 4. Дискретен канал без памет и без изтриване с ретранслация

Този модел обособява основните групи функции в отделните етапи на преобразуване на съобщенията и сигналите, както следва: ансамбълът от съобщения $\{u_i\} \rightarrow$ в множество от стойности на електрически сигнали $\{s(x_i)\}$; сигналите $\{s(x_i)\} \rightarrow$ в множество на информационните признаци $\{S(K^*)\}$; признаковото множество ($*$ – реализирани стойности) $\{S(K^*)\} \rightarrow$ в множество стойности на функцията за класификация $\{L[S(K^*)/D_i]\}$ (напр. ”функция на правдоподобие”, ”разстояние в някаква метрика” и т.н.); функцията за класификация $\{L[S(K^*)/D_i]\} \rightarrow$ в крайно решение $U(D_i)$ (електрически сигнал на входа на сортиращото устройство). Моделът от фиг.4 притежава всички преимущества на независимостта на анализите и оценките от голямото разнообразие, на описаните преобразувани сигнали, преимущества, осигурявани от приетия в теорията на информацията вероятностен подход.

В същото време е налице привидно противоречие с класическите модели по отношение състава на ансамблите на входа и изхода на канала. Действително, в създадения от нас модел, във връзка със специфичното

предназначение на пренасяната информация, се извършват поредица от целесъобразни преобразувания, което в крайна сметка довежда до редуциране на изходния ансамбъл спрямо входния. Ако се върнем към фиг.4 можем да опишем това редуциране по следния начин: нека приемем, че изходното множество $Z=\{z_1, z_2, \dots, z_j\}$, $j=1, 2, 3, \dots, M$ се разбива на подмножества $\{D_1, D_2, \dots, D_l, \dots, D_q\}$, $l=1, 2, 3, \dots, q$, при $q < M$. При това подмножествата $\{D_l\}$ са непресичащи се и всяко от тях включва от 1 до n съставни елементи z_j на множеството Z , а в съвкупността си те образуват множеството $D=\{D_1, D_2, \dots, D_l, \dots, D_q\}$ на изходния ансамбъл на канала от фиг. 4. Така имаме налице класическия дискретен канал без памет и без изтриване, с ретранслация $U, Z, P(Z/U)$ и създадения $U, D, P(D/U)$, за който е в сила неравенството $q < M$ за ансамблите на изхода и входа, съответно. Както беше показано, такъв канал няма физически смисъл като информационен канал в рамките на теорията на сигналите, което изисква да бъде обсъдена правомерността на третирането му наравно с класическите информационни канали. И на тази основа да бъдат използвани характеристиките и свойствата на тези канали при анализа на неговите функции. Ще докажем правомерността на такова разглеждане чрез следната теорема:

Ако множеството от съобщения $\{z_j\}$, $j=1, 2, 3, \dots, M$ на изхода на дискретен канал без памет и без изтриване се редуцира чрез целесъобразното им обединяване в нови, взаимно непресичащи се подмножества D_l , $l=1, 2, 3, \dots, q$ при $q < M$, това не може да доведе до увеличаване на реално получаваното на изхода количество информация I_p . За такъв канал могат да бъдат определени и условията за неразрушаване на пренасяната информация, т.е. след редуциране на изходния ансамбъл каналът запазва характера си на дискретен канал без памет и без изтриване (с ретранслация) в рамките на информационната теория.

Д о к а з а т е л с т в о: Нека запишем реално получаваното на изхода на канала количество информация (по Шенон) преди редуциране на изходното множество (7):

$$I_p(z) = H(U) - H(U/Z) = -\sum_{i=1}^M P(u_i) \log P(u_i) - \sum_{i=1}^M \sum_{j=1}^M P(z_j) [P(u_i/z_j) \log P(u_i/z_j)] \quad (8)$$

и след неговото редуциране

$$I_p(D) = H(U) - H(U/D) = -\sum_{i=1}^M P(u_i) \log P(u_i) - \sum_{i=1}^M \sum_{l=1}^q P(D_l) [P(u_i/D_l) \log P(u_i/D_l)] \quad (9)$$

По условие се иска да докажем, че $I_p(z) \geq I_p(D)$, т.е.

$$H(U) - H(U/Z) \geq H(U) - H(U/D), \quad (10)$$

за изпълнението на което е необходимо и достатъчно

$$H(U/Z) \leq H(U/D). \quad (11)$$

В теорията на информацията са известни съотношения, касаещи условните и съвместни ентропии на ансамбли от зависими съобщения:

$$1) H(U/Z) \leq H(U) + H(Z), \quad (12)$$

при това равенството се изпълнява тогава и само тогава, когато съобщенията z_j и u_i от двата ансамбъла са независими, т.е.

$$\sum_U \sum_Z P(u_i, z_j) \log P(u_i, z_j) \leq -\sum_U \sum_Z P(u_i, z_j) \log [P(u_i)P(z_j)] \quad (13)$$

само при $P(u_i, z_j) = P(u_i) \cdot P(z_j)$;

$$2) H(U, Z) = H(U) + H(Z/U) = H(Z) + H(U/Z), \quad (14)$$

където

$$H(U/Z) \leq H(U), \quad H(Z/U) \leq H(Z), \quad (15)$$

при това равенствата се изпълняват тогава и само тогава, когато U и Z са статистически независими, т.е.

$$-\sum_U \sum_Z P(u_i, z_i) \log P(u_i/z_i) \leq -\sum_U P(u_i) \log P(u_i) \quad (16)$$

само при $\sum_Z P(u_i, z_j) = P(u_i)$ и

$$\log P(u_i/z_j) = \log P(u_i), \text{ т.е.}$$

при статистически независими U и Z ;

$$3) H(A/B, C, D) \leq H(A/B, C) \quad (17)$$

при това равенството е възможно тогава и само тогава, когато

$$P(A, D/B, C) = P(A/B, C) P(D/B, C) \quad (18)$$

или, което е все едно

$$H(U/Z, D) \leq H(U/D), \quad (19)$$

при това равенството е възможно тогава и само тогава, когато

$$P(U, Z/D) = P(U/D) P(Z/D). \quad (20)$$

Нека разгледаме пространството на входния U и изходните Z и D ансамбли от съобщения. В това пространство може да бъде определено разпределение на вероятностите (множество на вероятностите, описващи връзките между случайните величини u_i , z_j и D_l) от вида $P(U, Z, D)$. Тъй като D_l са подмножества, съдържащи някакъв брой елементи на Z , то съвместните вероятности $P(U, Z, D)$ може да се запишат като равни на $P(U, Z)$, доколкото $\{z_j\} \in D_l$, $j=1, 2, 3, \dots$, или като равни на нула в противен случай. При това условие можем да изразим условната ентропия $H(U/Z)$ чрез съвместната вероятност $P(U, Z, D)$

$$H(U/Z) = -\sum_U \sum_Z \sum_D P(u_i, z_j, D_l) \log P(u_i/z_i, D_l) = H(U/Z, D) \quad (21)$$

От друга страна на основание на (19)

$$H(U/Z, D) \leq H(U/D),$$

следователно

$$H(U/Z) = H(U/Z, D) \leq H(U/D), \quad (22)$$

което съвпада с необходимото и достатъчно условие (11) за доказване на формулираното твърдение.

На базата на приведеното доказателство ще изследваме условията за неразрушаване на информацията в канала с редуциране на изходния ансамбъл. Съгласно (11), (19), (20), (22) неразрушаването (незагубването) на информацията в канала ще бъде възможно при удовлетворяване на равенството в

(11), (19), и (22), което е в сила при условие (20)

$$P(U, Z/D) = P(U/D) P(Z/D).$$

Когато подмножествата $\{z_j\}$, $j=1, 2, 3, \dots$, не принадлежат на $\{D_1\}$, т.е. $\{z_j\} \notin \{D_1\}$, това равенство се изпълнява безусловно. За нас интерес представлява случая на обединяване на съобщенията от Z в подмножества $\{z_j\}$, $j=1, 2, 3, \dots$ и принадлежността им (отнасянето априори) към някое от подмножествата $\{D_1\}$, т.е. $\{z_j\} \in D_1$, $j=1, 2, 3, \dots$. Ако запишем условните и съвместните вероятности, определени в пространството на входния и изходния ансамбли

$$P(U, Z/D) P(D) = P(U, Z, D), \quad (23)$$

където $P(U, Z, D)$ може да се запише и като $P(U, Z)$ по силата на $\{z_j\} \in D_1$, $j=1, 2, 3, \dots$ и

$$P(Z/D) P(D) = P(Z), \quad (24)$$

то търсеното условие с отчитане на (23) и (24) ще се свежда до равенството

$$P(U/Z) = P(U/D), \text{ при } \{z_j\} \in D_1, j=1, 2, \dots \quad (25)$$

в областите на пространството $\{U, Z, D\}$, където лявата и дясната му страна са определени.

ЗАКЛЮЧЕНИЕ

Изследвани са теоретични подходи за усъвършенстване на системите за обектив-

но (машинно) определяне на качеството на аграрни продукти и други храни.

Установени са възможности за адаптиране на добре развития теоретичен апарат на общата теория на сигналите и теорията на информацията към задачите по създаване на нови технически средства и системи с автоматични класификатори по качество на храни. Трансформирането на тези системи в информационни канали позволява откриване от конкретната им физическа същност, моделиране и изследване в условията на голямото им разнообразие, което очертава перспективността на подхода.

ЛИТЕРАТУРА

- [1] Feinstein A., Foundations of information theory, McGraw-hill Inc. New York, 1958
- [2] Georgiev A., L. Kostadinova and R. Gabrova, "Problems of discrimination and quantization of optic-electronic images of agricultural products in automatic quality classifiers" Scientific Works UFT, vol.XLIV, Iss. 3, pp.286-296.
- [3] Иларионов Р., Компютърна периферия, УИ Васил Априлов, Габрово, 2008
- [4] Иларионов, Р., И. Симеонов, Н. Шопов, Х. Килифарев. Компютърно базирана система за ултразвукова идентификация на хранителни продукти. Сборник доклади от Международна научна конференция УНИТЕХ'09, Том I, 20-21 ноември 2009 Габрово, стр. I-212 – I-218.