

Оценка влиянието на сортовата принадлежност върху цветови признаци за разпознаване на заболяването Фузариоза по царевични семена

Е. Кирилова, П. Даскалов, Р. Цонев, Ц. Драганова

Evaluation of the influence of different classes of corn kernels in reference to the recognition of *Fusarium*- damaged seeds by color features: Statistical evaluation of the color features for 8 classes of corn kernels, that are related to identification of *Fusarium*- damaged corn seeds by digital image analysis is presented in the paper. The aim of this study was to determine whether the image parameters of different classes of corn seeds were statistically different from each other or not. The color features were compared by applying one way ANOVA (analysis of variance) method followed by Tukey's post hoc test using the software STATISTICA 8, with a 0.05 significance level.

Key words Statistical approaches, One way ANOVA, corn kernels, Image processing.

ВЪВЕДЕНИЕ

Основната тенденция за бързо и обективно окачествяване на семена и зърнени култури налага използване на компютърното зрение при идентифицирането и анализа на продукцията [3]. Стремешът да се разработят универсални алгоритми за класификация е свързан с провеждане на задълбочен анализ на зависимостите между основните характеристики на изображенията на семена на различни култури от различни сортове [1], [4], [5], [7], [8] както и на други хранителни продукти [6], [9].

Имайки предвид изложеното дотук, основна цел на настоящия доклад е да се оцени влиянието на сортовата принадлежност върху цветови признаци за разпознаване на заболяването *Fusarium Moniliforme* по царевични семена с помощта на статистически методи, като се анализират съществуващи зависимости между статистическите оценки на различни цветови признаци за осем сорта семена.

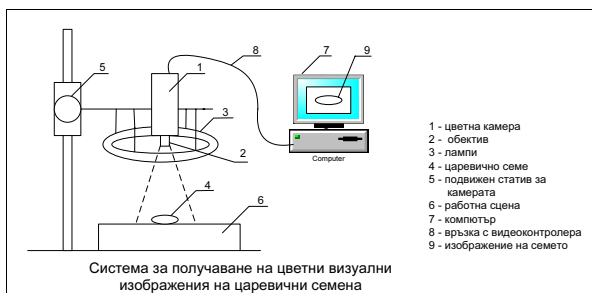
ИЗЛОЖЕНИЕ

1.Обект и методика

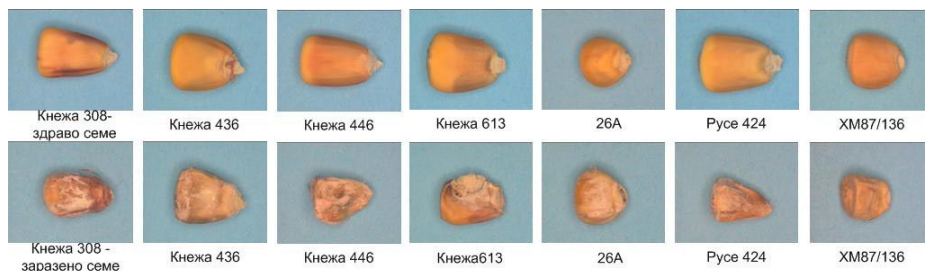
Обект на изследването са здрави и заразени от розова фузариоза *Fusarium Moniliforme* семена от 8 сорта. Обема на анализирани извадки семена включва съответно по 150 здрави и 150 болни семена от сортове 26А, Кнежа 308, Кнежа 436, Кнежа 446, Кнежа 613, Кнежа 620, 90 здрави и 90 болни семена от сорт Русе 424 и 60 здрави и 60 болни семена от сорт ХМ87/136.

Изображенията на семената са получени с размери (352 x 289) пиксела, с 8 битово кодиране за червената (R) , зелената (G) и синята (B) компоненти на RGB цветовия модел. Използвана е цветна камера: PIH – 7030 H / IR с разделителна способност 1/3" CCD, 752 x 582 pix (352 x 289), 460 TV линии, 16 млн. цвята и обектив Tokina, който позволява ръчна промяна на фокусното разстояние от 6 до 60 mm. Разстоянието между камерата и семето при снемане на изображенията е 26 cm. Осветлението е осигурено от 2 бр. кръгли луминисцентни лампи над работната сцена. На фиг.1. е показана схема на системата за снемане на изображения.

По нататъшната обработка на изображенията включва сегментация по H (Hue) компонентата на HSV цветовия модел, като целта е да се отделят пикселите принадлежащи на фона, от тези, принадлежащи на обекта. Пикселите, принадлежащи на семето се преобразувани в HSV, Lab, XYZ, YCbCr, xyY цветови модели. На фиг.2 се виждат изображения на здрави и заразени царевични семена от осемте сорта.



Фиг.1. Схема на системата за получаване на цветни цифрови изображения



Фиг.2. Цветни цифрови изображения на царевични семена

За целите на анализа е необходимо е да се сравнят стойностите на признаците (компонентите на цветовите модели) за семената от различните сортове, отчитайки състоянието на семето - здраво или заражено. Групите здрави и заразени семена от различните сортове разглеждаме като независими извадки, т.е. разполагаме с 16 извадки с различен обем, касаещи всеки от 17-те признака.

Базите данни съдържат средноаритметични стойности на 17 компоненти от 6 цветови модела, както следва: R, G, B, L, a, b, X, Y, Z, H, S, V, Y, Cb, Cr, x, y за здрави и заразени семена от осем сорта.

$$\bar{R}_{t,s}(RGB) = \frac{1}{n_t} \sum_{i=1}^{n_t} R_{ti} \quad (1)$$

$\bar{R}_{t,s}(RGB)$ е средноаритметичната стойност на R компонентата (от RGB цветови модел) за пикселите от изображението на t-то семе от сорт s;

$t=1,2,\dots,n_s$; n_t - брой на пикселите $\in$ на изображението на t-то семе;

R_{ti} - стойност на R компонентата в i-ия пиксел на t-то семе.

Необходимо е да бъде проверена хипотезата за равенство на средните стойности на цветовите признаци за здравите семена от осемте сорта, а също и за болните семена от осемте сорта, като са извършени два независими анализа - за здрави и за заразени семена. Разполага се с $m=8$ независими извадки с обеми $n_1, n_2, \dots, n_m$, които принадлежат на нормално разпределени генерални съвкупности. Искане се да се провери хипотезата $H_0: E[X_1] = E[X_2] = \dots = E[X_m]$, където $E[X_j], j=1,2,\dots,m$ са средните стойности на семената от съответните

сортове за даден цетови признак. Проверката на хипотези от този вид изразява същността на еднофакторния дисперсионен анализ.

Експерименталните данни са анализирани с програмния пакет на StatSoft STATISTICA 8. Приложен е еднофакторен дисперсионен анализ (one-way ANOVA). Ако стойността на степента на вероятност p - level (Statistical Significance level) е по-малка от $\alpha=0.05$, то нулевата хипотеза H_0 се отхвърля и се счита, че поне за една от групите средната стойност е значително различна от средните стойности на останалите групи. За да бъдат определени конкретно групите, чиито средни стойности са статистически различни е необходимо да се проведат следващи тестове, известни като тестове за множествени сравнения (Post-hoc tests) [2]. За целите на настоящия анализ е удачно да се проведе тест за множествено сравнение по метода на Тюки (Tukey HSD test). При тази процедура се извършват сравнения между всички двойки средни стойности, като се запазва предварително зададена стойност за грешката. Тестовете от модула ANOVA се изпълняват при характеризирането на всяка от категориите семена, като се търсят значими различия между средните стойности за всеки клас. При необходимост да бъдат сравнени само две средни стойности от 2 различни групи наблюдения, то ANOVA дава същите резултати както и Т-теста за независими извадки.

2. Получени резултати и анализ

Резултати от еднофакторен дисперсионен анализ при ниво на значимост $\alpha=0.05$ са дадени в таблица 1. При анализа независимата величина е факторът сорт с осем поднива (сорт1,сорт2,...,сорт8), а зависимите величини са 17-те цетови признака.

Таблица 1. Резултати от статистическа процедура ANOVA за 8 сорта семена, за 17 цетови признаци

Признак	F- здрави семена	F- заразени семена
Z_mean	130,2740	40,8944
Xmean	105,5233	26,0880
Ymean	106,0475	23,6363
Rmean	65,4738	42,2823
Gmean	130,4019	22,3526
Bmean	126,2759	44,6194
Lmean	122,7089	25,1460
amean	153,2873	18,8616
bmean	97,9377	45,5262
Hmean	106,7233	130,6678
Smean	61,9238	63,1948
Vmean	65,3201	38,8643
Ycbcrmean	115,4733	24,3132
cbmean	72,5722	60,9336
crmean	73,7835	50,8220
xmmean	89,2300	34,3868
ymmean	142,5728	49,8431

В табл. 1 са показани резултатите от еднофакторен дисперсионен анализ върху данните за 17-те компоненти на изследваните цетови модели извлечени от изображенията за здрави и заразени семена. Степента на вероятност p -level за всички признаци има нулеви стойности, т.е стойностите са по-малки от критичната

$p=0.05$, поради което нулевата хипотеза H_0 се отхвърля и се счита, че поне за една от групите средната стойност е значително различна от средните стойности на останалите групи. Резултатите от анализа one-way ANOVA показват, че осемте сорта се различават статистически по всяка една от зависимите величини.

Получените резултати, които отхвърлят нулевата хипотеза за равенство на средните стойности на групите не уточняват между кои средни се наблюдава достоверна разлика, това означава, че средните на групите трябва да се сравняват по двойки с помощта на тестовите за множествени сравнения.

При анализа имената на сортовете са кодирани както следва:

S1 – сорт Кнежа 446, здрави семена	D1 – сорт Кнежа 446, заразени семена
S2 - Кнежа 613, здрави семена	D2 - Кнежа 613, заразени семена
S3 – Кнежа 620, здрави семена	D3 – Кнежа 620, заразени семена
S4 – 26A, здрави семена	D4 – 26A, заразени семена
S5 – Кнежа 308, здрави семена	D5 – Кнежа 308, заразени семена
S6 – Кнежа 436, здрави семена	D6 – Кнежа 436, заразени семена
S7 – Русе 424, здрави семена	D7 – Русе 424, заразени семена
S8 – XM87/136, здрави семена	D8 – XM87/136, заразени семена

Таблица 3. Обобщени резултати от тест на Тюки за здрави семена

сорт	RGB			HSV			Lab			XYZ			YCbCr			xyY	
	R	G	B	H	S	V	L	a	b	X	Y	Z	Y	cb	cr	x	y
S1 - S2	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
S1 - S3	-	-	-	-	+	+	-	+	-	-	-	-	+	+	+	-	+
S1 - S4	-	-	+	-	+	-	-	+	+	-	-	-	-	+	+	+	+
S1 - S5	-	+	-	-	-	-	+	-	-	+	+	-	+	-	-	-	-
S1 - S6	-	+	-	-	-	-	+	+	-	+	+	-	+	-	-	-	-
S1 - S7	-	-	+	-	+	-	+	+	-	-	+	-	+	+	+	+	+
S1 - S8	+	-	+	+	+	+	-	-	+	-	-	+	-	+	+	+	+
S2 - S3	-	-	-	+	-	+	-	-	+	-	-	-	-	+	+	-	+
S2 - S4	-	-	-	+	+	-	-	-	+	-	-	-	-	+	+	+	+
S2 - S5	-	-	-	-	-	+	-	-	+	+	-	+	+	-	-	-	-
S2 - S6	-	+	+	+	+	-	+	+	-	+	+	-	+	-	+	+	-
S2 - S7	-	-	+	+	+	-	-	+	+	-	-	+	-	+	+	+	+
S2 - S8	+	-	-	-	+	+	-	-	+	-	-	+	-	+	+	+	+
S3 - S4	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
S3 - S5	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
S3 - S6	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-
S3 - S7	-	-	-	+	+	-	-	-	+	-	-	-	-	+	+	+	+
S3 - S8	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
S4 - S5	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
S4 - S6	-	-	-	+	-	+	-	+	-	-	-	-	-	-	+	-	-
S4 - S7	-	-	-	+	+	-	-	+	+	-	-	-	-	-	+	+	+
S4 - S8	-	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
S5 - S6	-	+	-	-	-	+	-	-	-	+	-	+	-	-	-	-	-
S5 - S7	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
S5 - S8	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
S6 - S7	+	-	+	+	-	-	+	-	-	-	-	+	-	-	-	-	-
S6 - S8	-	-	-	-	+	+	-	-	+	+	+	-	+	-	-	-	-
S7 - S8	-	-	-	-	+	-	-	-	+	-	-	-	+	+	+	+	+

В табл. 3 и 4 са дадени обобщените резултати от теста за множествено сравнение по метода на Тюки, като със знак + е отбелязано равенство на средните стойности на цетовите признаци на съответните сортове, а знак – означава, че има статистическа разлика между съответните средни стойности.

Таблица 4. Обобщени резултати от тест на Тюки за заразени семена

сорт	RGB			HSV			Lab			XYZ			YCbCr			xyY	
	R	G	B	H	S	V	L	a	b	X	Y	Z	Y	cb	cr	x	y
D1 – D2	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
D1 – D3	-	-	-	-	+	+	-	-	+	-	-	-	-	+	+	-	+
D1 – D4	-	-	+	-	+	-	-	+	+	-	-	-	-	+	+	+	+
D1 – D5	-	+	-	-	-	-	+	-	+	+	+	+	-	-	-	-	-
D1 – D6	-	+	-	-	-	-	+	+	-	+	+	+	+	-	-	-	-
D1 – D7	-	-	+	-	+	-	-	+	+	-	-	+	-	+	+	+	+
D1 – D8	+	-	+	+	+	+	-	-	+	-	-	+	-	+	+	+	+
D2 – D3	+	-	-	+	-	+	-	+	-	-	-	-	-	+	+	-	+
D2 – D4	-	-	-	+	+	-	-	+	-	-	-	-	-	+	+	+	+
D2 – D5	-	-	-	-	-	-	+	-	-	+	+	-	+	-	-	-	-
D2 – D6	-	+	+	+	+	-	+	+	-	+	+	+	+	-	+	+	-
D2 – D7	-	-	+	+	+	-	+	+	-	-	+	-	-	+	+	+	+
D2 – D8	+	-	-	-	+	+	-	-	+	-	-	-	-	+	+	+	+
D3 – D4	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
D3 – D5	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D3 – D6	+	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-
D3 – D7	-	-	-	+	+	-	-	+	-	-	-	-	-	+	+	+	+
D3 – D8	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
D4 – D5	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D4 – D6	+	-	-	+	-	+	-	+	-	-	-	-	-	-	+	-	-
D4 – D7	-	-	-	+	+	-	-	+	+	-	-	-	-	-	+	+	+
D4 – D8	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
D5 – D6	-	+	-	-	-	-	+	-	-	-	+	-	+	-	-	-	-
D5 – D7	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-
D5 – D8	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
D6 – D7	-	-	+	+	-	-	+	-	-	-	+	-	-	-	-	-	-
D6 – D8	+	-	-	-	-	+	+	-	-	+	+	-	+	-	-	-	-
D7 – D8	-	-	-	-	+	-	-	-	+	-	-	-	-	+	+	+	+

За да бъдат сравнени две по две средните стойности между здравите и заразените извадки семена от един и същ сорт е изпълнен Т-тест на Стюдънт за независими извадки. Резултатите от проведения тест са дадени в таблица 5.

Таблица 5. Резултати от Т-тест за независими извадки

сорт	RGB			HSV			Lab			XYZ			YCbCr			xyY	
	R	G	B	H	S	V	L	a	b	X	Y	Z	Y	cb	cr	x	y
Кнежа 446 S1 – D1	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-
Кнежа 613 S2 – D2	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
Кнежа 620 S3 – D3	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
26 A S4 – D4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Кнежа 308 S5 – D5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Кнежа 436 S6 – D6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Русе 424 S7 – D7	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-
ХМ87/136 S8 – D8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

ЗАКЛЮЧЕНИЕ

Приложеният еднофакторен дисперсионен анализ показва, че факторът сорт има статистически значимо влияние върху средните стойности на всички 17 цветовите признаци, както за здравите, така и за заразените семена. Това е показател, че налепът по царевичните семена, който се образува вследствие на заболяването Фузариоза има различни цветови характеристики за различните сортове. Резултатите от проведения тест на Тюки показват, че сортове Кнежа 308 и Кнежа 436 се различават по изследваните 17 цветови характеристики от останалите сортове, а имат статистически равни средни стойности на следните признаци -

L(Lab), X, Y(XYZ) и Y(YCbCr). Сортове Кнежа 446, Кнежа 613, Кнежа 620, 26А, Русе 424 и ХМ87/136 имат статистически равни средни стойности за други три признака, но от същите цветови модели – b(Lab), Cr(YCbCr) и y(xyY).

За определяне на информативни признаци, които да бъдат използвани при разпознаването на заболяването Фузариоза, е направена проверка на хипотезата за равенство на средните им стойности между здрави и заразени семена от един и същ сорт. Резултатите показват, че за сортове Кнежа 446, Кнежа 613, Кнежа 620 и Русе 424 средните стойности, които са статистически равни са съответно на следните признаци: L(Lab), Y(XYZ) и Y(YCbCr) за Кнежа 446; G(RGB) и H(HSV) за Кнежа 613; H(HSV) за Кнежа 620; Z(XYZ) за Кнежа 620. При разпознаване на Фузариоза по царевични семена е необходимо да бъдат използвани различни групи признаци за отделните сортове, за които няма статистическо равенство между стойностите за здрави и заразени семена.

Изследванията са подкрепени по договор № **BG051PO001-3.3.04/28**, „Подкрепа за развитие на научните кадри в областта на инженерните научни изследвания и иновациите“. **Проектът се осъществява с финансовата подкрепа на Оперативна програма „Развитие на човешките ресурси“ 2007-2013, съфинансирана от Европейския социален фонд на Европейския съюз“.**

ЛИТЕРАТУРА

- [1] Choundary R., J. Paliwal, D.S. Jayas Classification of cereal grains using wavelet, morphological, colour and textural features of non-touching kernel images, Biosystem Engineering 99 (2008) 330-337
- [2] Brown Angus M. A new software for carrying out one-way ANOVA post hoc tests, Computer Methods and Programs in Biomedicine (2005) 79, 89—95
- [3] Brosnan Tadhg, Da-Wen Sun Inspection and grading of agricultural and food products by computer vision systems – a review, Computers and Electronics in Agriculture 36 (2002), 193-213
- [4] Dana Wiesnerova, Ivo Wiesner Computer image analysis of seed shape and seed color for flax cultivar descriptor, Computers and Electronics in Agriculture 61(2008),126-135
- [5] Granitto Pablo M., H.D. Navone, P.F. Verdes, H.A. Ceccatto Weed seeds identification by machine vision, Computers and Electronics in Agriculture 33(2002), 91-103
- [6] Kavdir I., D.E. Guyer Evaluation of different pattern recognition techniques for apple sorting, Biosystem Engineering 99 (2008) 211-219
- [7] Majumdar S., D.S. Jayas Classification of bulk samples of cereal grains using machine vision, J.Agric. Engng. Res. (1999) 73, 35-47
- [8] Manickavasagan A., G. Sathya, D.S. Jayas, N.D.G. White Wheat class identification using monochrome images, Journal of Cereal Science 47 (2008), 518-527
- [9] Riquelme M.T., P. Barreiro, M. Ruiz-Altisent, C. Valero Olive classification according to external damage using image analysis, Journal of Food Engineering 87 (2008) 371–379

За контакти:

инж. Елеонора Стефанова Кирилова, Катедра “Автоматика, информационна и управляваща техника”, Русенски Университет “Ангел Кънчев”, тел.: 082/888 684, e-mail : ekirilova@ru.acad.bg

Доц. д-р инж. Пламен Даскалов, Катедра “Автоматика, Информационна и Управляваща Техника”, Русенски университет “Ангел Кънчев”, тел.:082 888668, E-mail: daskalov@ru.acad.bg

гл.ас.д-р Цветелина Драганова, Катедра “Автоматика, Информационна и Управляваща Техника”, Русенски университет “Ангел Кънчев”, тел.:082 888668, E-mail: cgeorgieva@ru.acad.bg

Докладът е рецензиран.